Evaluation of the IBL Mathematics Project: Student and Instructor Outcomes of Inquiry-Based Learning in College Mathematics

A Report Prepared for the Educational Advancement Foundation

and the IBL Mathematics Centers

Assessment & Evaluation Center for Inquiry-Based Learning in Mathematics

Sandra Laursen, Marja-Liisa Hassi, Marina Kogan, Anne-Barrie Hunter

Ethnography & Evaluation Research

and Tim Weston

ATLAS Assessment and Research Center

University of Colorado Boulder

April 2011

Executive Summary

Evaluation of the IBL Mathematics Project: Student and Instructor Outcomes of Inquiry-Based Learning in College Mathematics

Assessment & Evaluation Center for Inquiry-Based Learning in Mathematics

Sandra Laursen, Marja-Liisa Hassi, Marina Kogan, Anne-Barrie Hunter,

Ethnography & Evaluation Research,

and Tim Weston, ATLAS Assessment and Research Center

University of Colorado Boulder

April 2011

ES 1.0 Introduction

We report findings from a comprehensive, mixed-methods study of inquiry-based learning (IBL) in college mathematics as implemented at four university IBL Mathematics Centers. Six sub-studies were collectively designed to examine the following questions:

- What are the student outcomes of IBL mathematics courses?
- How do these outcomes vary among student groups, and how do they compare with other types of courses?
- How do these outcomes come about?
- What are the costs and benefits of IBL teaching for instructors and departments?

Data analyzed for this report include approximately 300 hours of classroom observation, 1100 surveys, 220 tests, 3200 student transcripts, and 110 interviews with students, faculty, and graduate teaching assistants (TAs), gathered from over 100 course sections at four campuses across two academic years.

In the main report,[1] we introduce the IBL Mathematics Centers, the study and its context (Chapter 1), then document findings from the six sub-studies (Chapters 2-8). Our summary (Chapter 9) gathers key findings, emphasizing corroborating evidence from multiple lines of inquiry. The appendices (A1-A7) document the study methods and samples for each sub-study. In the summary below, section numbers are provided to reference discussion of specific findings.

---

[1] http://www.colorado.edu/eer/research/steminquiry.html#Reports

ES 2.0 Setting for the Study

The study sites were four university mathematics departments that host IBL Mathematics Centers supported by the Educational Advancement Foundation. All four Centers are at research universities (three public, one private) with selective undergraduate enrollment and highly ranked graduate programs in mathematics. They are geographically dispersed and vary in size. Because the IBL Mathematics Centers were established in 2004, before this study was commissioned, many study parameters were pre-determined. The study used an observational, not experimental, design, reflecting realistically variable implementations of inquiry-based learning and not an idealized laboratory situation. That circumstance imposed some constraints on what we could learn, but also makes our findings relevant to other real-world reform efforts in STEM higher education.

ES 2.1 Courses Studied

Starting in 2004, each Center independently chose certain courses in which to develop and teach IBL mathematics curricula. These courses were well established by the time our study began, and in most cases multiple instructors had participated in addition to the course originators. Varying levels of professional development and support were available to instructors new to IBL. Most commonly, they visited colleagues' classes, drew upon others' course materials, and exchanged ideas informally or at occasional IBL events.

The courses included in the study addressed a range of student audiences and mathematical topics. Courses for pre-service K-12 teachers focused on deep understanding of the mathematical concepts needed to teach elementary, middle, or secondary schoolchildren. Courses such as cryptology and calculus sought to provide talented first-year students with a stimulating mathematics experience that might draw them into the major. More advanced courses in number theory, analysis, geometry, and discrete math served "mainstream" mathematics majors as well as science and engineering majors.

Some IBL courses were offered with parallel non-IBL sections also available. Others—including all of the courses targeted to pre-service K-12 teachers—were offered only in IBL format. Thus the availability of comparison groups for the study was uneven. In the analyses discussed below, data from first-year and advanced courses were combined in describing "math-track" courses, while data from courses for pre-service teachers were analyzed separately.

ES 2.2 Inquiry-Based Learning as Implemented at the IBL Centers

We labeled course sections as "IBL" or "non-IBL" following the campus Centers' own designations, which generally indicated support by the Center's grant. To establish that the IBL label indeed differentiated teaching and learning approaches, we observed classrooms to directly establish the nature of instruction actually occurring in IBL and non-IBL courses.

Students in IBL classes spent class time giving and listening to student presentations, working in small groups, discussing ideas that generally arose from a group problem or student-presented solution, and, in a few courses, working with peers on computer activities such as modeling or simulations (2.2.1). On average, over 60% of IBL class time was spent on such student-centered activities, while students in non-IBL courses spent 87% of their time listening to their instructor talk. While there was substantial variation in practice, overall the IBL courses offered students very different experiences from typical lecture-based courses. That is, as a group, instruction was clearly distinct between "reformed" IBL courses and comparative non-IBL sections.

Drawing on both the classroom observation data (Ch. 2) and student and instructor interviews (7.3; 8.2.6, 8.3), we identify key features of IBL courses in this study by both their statistical frequency and their importance to students and instructors. In IBL courses typical of this project:

- The main work of the course was problem-solving: students solved challenging problems alone or in groups, in and out of class. In class they shared, evaluated and refined their own and each others' solutions.
- The course was driven by a carefully built sequence of problems or proofs, rather than a textbook. The pace of the course was set by students' movement through this sequence, rather than pegged to a pre-set schedule.
- Course goals tended to emphasize development of skills such as problem-solving, communication, and mathematical habits of mind, not just covering specific content.
- Most of class time was spent on student-centered instructional activities. Students or groups of students played a leading role in guiding these activities. Most activities were conducted for just a few minutes at a time: class work tended to change gears often and switched frequently between activities.
- Instructors' main role was not lecturing. They might give mini-lectures to set up or sum up the day's work, introduce a group activity, or provide context for a set of theorems. Instructors (as well as other students) may offer impromptu explanations to respond to a comment or question. They might ask questions to clarify student thinking, refine a proposed solution, give feedback, or elicit such comments from other students.
- Student voices were heard in the classroom: presenting, explaining, arguing, asking questions. Their active participation meant that students themselves had considerable influence on how class time was spent and how fast the class moved through the material.
- This joint responsibility for the depth and progress of the course fostered a collegial atmosphere that placed value on respectful listening and critique and that invited every class member to contribute fruitfully to the mutual development of mathematical ideas. Instructors made efforts to set and maintain this atmosphere.

From instructor[2] and student interviews, we also know something about the behind-the-scenes structure typical of an IBL course. Outside class, much of students' work time—which was

[2] The term instructor refers to faculty and TAs together.

substantial—was spent in preparing for class: solving problems or deriving proofs to present or discuss in groups. Because work was due nearly every class, the workload was steady rather than test-driven (3.3.2). Instructors invested substantial time in constructing the "script," or sequence of problems, or in understanding and fine-tuning scripts shared by other instructors. Checking homework took on greater importance for IBL courses, because students' work improved most rapidly when they got timely feedback on their solutions or proofs. Faculty and TAs held many office hours outside class, and students made much use of them, because timely help could be important to making progress and separating "fruitful struggle" from wasted time.

ES 3.0 Student Learning Outcomes

Findings on student learning outcomes were derived from several independent lines of evidence:

- surveys of students' self-assessed learning gains (Ch. 3), which include both numerical ratings and open-ended comments;
- tests given to subsets of math-track and pre-service students (Ch. 5);
- interviews with students (Ch. 7) and their instructors (Ch. 8); and,
- for math-track students, analysis of their grades and course-taking patterns subsequent to an IBL or comparative course (Ch. 6).

On surveys of their self-assessed learning gains, math-track students who took IBL courses reported greater gains than their peers who took non-IBL sections of the same courses (3.2.1). These gains were higher across several areas: cognitive gains, including understanding of mathematical concepts and improved thinking and problem-solving skills; affective gains, including confidence, positive attitude, and persistence; and social gains, including collaboration and comfort in teaching mathematical ideas to others.

For pre-service teachers who took IBL courses, the self-reported learning gains were different (3.2.1). In general, their cognitive and affective gains were lower than those of math-track IBL students. However, their gains in applying mathematical knowledge, collaboration, and comfort in teaching mathematical ideas to others were as strong or stronger than those of math-track IBL students, and clearly higher than those of math-track non-IBL students. These differences likely reflect pre-service teachers' lower general interest in mathematics (3.4.3), but also emphasize their gains in areas that are especially significant for their future work in teaching K-12 students.

Students' write-in comments on surveys reinforce their survey ratings about the breadth and depth of learning they experienced (3.2.2). From both math-track and pre-service courses, twice as many IBL students wrote about learning gains as did non-IBL math-track students, and IBL students also wrote much more often of multiple gains from their course. Their comments emphasized cognitive gains—especially learning more deeply because they had figured things out for themselves. They also described changes in how they learned mathematics and solved problems, including improved learning on their own and from others. They described affective benefits and new communication skills—less often than they noted cognitive and learning changes, but more often than either type was noted by non-IBL students.

Student interview data further corroborate these findings. Again, students most emphasized cognitive gains (7.2.1), especially the deep and lasting learning that came from working through ideas for themselves. They saw their gains in thinking and problem-solving skills as transferable to other courses and to life in general. Changes in learning, also commonly reported, were of two types: personal learning changes such as self-awareness, persistence and independence, and greater appreciation for the benefits of collaborative work (7.2.3). Affective gains emphasized confidence, enjoyment and interest (7.2.4). They noted communication gains too: improved writing and speaking about proofs and enhanced abilities to explain and critique ideas (7.2.5).

Instructor observations of student learning aligned well with students' own reports, with some differences in perspective (8.2.1-8.2.4). Instructors could better spot gains in communication skills and understanding the nature of mathematics that indicated students' development as budding mathematicians, but less easily observed gains in students' personal learning processes.

For pre-service teachers, pre- to post-test score gains on a well-validated external test of learning mathematics for teaching (LMT) reflected real gains in understanding after an IBL course (5.2.2). In prior work, improved scores on this test have been connected to positive effects on teachers' instruction. The LMT test results thus suggest that IBL courses benefited pre-service teaching students in ways that will benefit their future work as teachers.

With a sample of math-track students, we gave a "proof test" to compare IBL and non-IBL students' ability to evaluate mathematical arguments and their reasons for judging arguments to be proofs or not. Both groups performed well on the test; there were no general differences in their scores overall or on specific problems (5.3.3). However, there was some evidence that IBL students were more skilled in recognizing valid and invalid arguments (5.3.3) and that they applied more expert-like reasoning in making such evaluations (5.3.6). The small sample size of academically strong students limits interpretation of these findings, as effects of IBL courses on performance may be greater for students whose prior academic record is less strong (see 4.2).

Academic records data from three target courses indicated that IBL students earned grades in subsequent courses that were as good or better than the grades of their non-IBL peers (6.2.1). The pattern favoring IBL students was broadly consistent across different subsets of later courses, but few of the differences were statistically significant, due to the wide scatter in the grades of both groups. Overall, taking IBL courses may benefit, and certainly does not harm, students' performance in later mathematics courses.

Among students who took mid-level and advanced courses, there were no general differences in pursuit of additional mathematics courses (6.2.2). However, students who took an IBL honors course early in their college career did appear to take more subsequent math courses than did a matched sample of peers from the large lecture-based section. Most of these differences are suggestive rather than definitive because of small sample sizes, but the IBL students did complete further IBL courses at a statistically much higher rate. These results imply that IBL honors courses can draw strong students into further study, especially additional IBL courses. Later IBL experiences appeared to neither spur nor deter further study of mathematics.

Overall, several lines of evidence suggest that students who had a college IBL course grew as mathematicians and as learners in ways that their peers taking non-IBL courses typically did not. The nature and types of the observed cognitive, affective and social gains were very consistent across multiple data sources. Some of IBL students' cognitive gains in reasoning and problem-solving were detectable on tests; and there is some evidence that these gains carry over to benefit students' work in later courses. Of all the learning outcome measures that we compared between IBL and non-IBL students, very few pointed to deficits for IBL students.

ES 4.0 Group Differences in Student Learning Outcomes

Our findings on group differences in student learning outcomes are based on subdividing the same data sets listed in ES 3.0: student surveys, tests, interviews, and academic records.

ES 4.1 Group Differences by Gender

Women's share of undergraduate degrees in mathematics has declined in the past two decades, unlike most other STEM fields. Thus we examined differences in student outcomes by gender.

In survey items on self-assessed learning gains, women in IBL classes reported as high or higher gains than their male classmates across all cognitive, affective and social gains areas (3.2.3). But women in non-IBL classes reported statistically much lower gains than their male classmates in several important domains: understanding concepts, thinking and problem-solving, confidence, and positive attitude toward mathematics. Overall, both men and women reported higher learning gains from IBL courses than from non-IBL courses, while traditional teaching approaches did disservice to women in particular, inhibiting their learning and reducing their confidence. Women's spontaneous write-in comments echoed this finding: IBL women wrote over four times as many comments about their cognitive gains, in particular, compared to their non-IBL peers, and also more comments about gains in confidence. IBL approaches appeared to level the playing field for women, compared to traditional lecture-based approaches.

The academic records data offer perspective on how well these short-term gains may last for women in later courses. IBL women outperformed their non-IBL counterparts on several measures of subsequent grades (6.3.1), as did IBL men (6.3.2). IBL women also took as many or more courses than non-IBL women (6.5.1), though generally fewer than their male classmates (Figures 6.7-6.9). These patterns were fairly consistent but mostly not statistically significant.

How women's grades compared to their male classmates seemed to depend somewhat on the course level (6.3.3.). Women in the most advanced course held their ground versus their male classmates, but women in the first-year course underperformed both their female non-IBL peers and their male IBL classmates. In the mid-level course, results for women fell roughly between the other two. First-year IBL women also pursued further IBL courses at lower rates than their male classmates. Perhaps women taking advanced courses are survivors who have learned to thrive independent of the challenges they encounter, while women early in their college careers may be more sensitive to stereotype threat and other differences in real or perceived classroom climate. While the short-term results indicate that IBL experiences particularly benefit women, over the long haul, a single IBL experience may help to close but not erase the gender gap.

There were no gender differences in the gains reported by the students we interviewed (Table 7.2, 7.4), and very few reports of gender-based differences in their experiences. The fact that gender is such a non-issue in the interview data explains the survey data rather well: from students' perspective, IBL classrooms offered equitable environments where all could succeed.

Overall, IBL experiences appear to be powerful for women, leveling the playing field by eliminating discouraging experiences that impede learning in traditionally taught courses.

ES 4.2 Group Differences by Prior Achievement

Instructors commonly hypothesized that IBL experiences were most beneficial for students who were not the most academically qualified—students who were good, perhaps, but not great (8.2.5). Thus we examined the data for differences in student outcomes by prior achievement.[3]

On the LMT test given to IBL pre-service teachers, test score gains were anti-correlated with the initial score (5.2.3). That is, students who had the lowest scores on the pre-test improved the most on the post-test. This finding matches students' self-report of learning gains, where IBL pre-service teachers with lower overall GPAs reported higher gains than their classmates with higher grades (3.2.6), mirroring the pattern seen on the LMT test.

For math-track students, the pattern among IBL students was less striking, but similar. Students with the lowest GPAs reported higher gains than the middle GPA group (3.2.6). However, in non-IBL classes, students who had the highest GPAs reported the highest gains, and low-GPA students reported the lowest gains.

Among students who entered with low math GPAs (<2.5), IBL students generally earned better grades in later classes than did their non-IBL peers (6.4.1). IBL students who entered with higher math grades also did as well or better than non-IBL peers in later courses (6.4.2, 6.4.3), but the improvement for previously low-achieving IBL students was striking (6.4.4.).

Overall, it appeared that non-IBL courses tended to reinforce prior achievement patterns, helping the "rich" to get "richer." In contrast, IBL courses seemed to offer an extra boost to lower-achieving students, especially among pre-service teachers. Yet there was no evidence of harm done to the strongest students. Indeed, high-achieving students may be encouraged by an IBL experience to take more mathematics courses, especially more IBL courses (6.6.2)—again, consistent with instructor observations that strong students found the IBL approach stimulating

---

[3] Different types of data were necessarily used to distinguish prior achievement in each sub-study.

(8.2.5). Instructors did have reservations about the benefits of IBL for the "weakest" students; and our analyses may not have detected a small subset of students who genuinely struggled.

ES 4.3 Group Differences by Experience Level

In the pilot study, first-year students were particularly enthusiastic about how IBL courses had enhanced their learning in other courses. Thus we investigated differences in student outcomes for students taking IBL courses earlier or later in their college careers.

On surveys, first-year math-track students who took IBL courses reported higher gains than did IBL students later in their careers, across several areas: mathematical thinking, persistence, and collaboration (3.2.4). Moreover, both first-year and mid-career (sophomore/junior) students reported higher gains in these areas, as well as confidence and positive attitude about mathematics, compared to their non-IBL peers. Even late-stage (senior and graduate) students reported higher affective and collaboration gains than did their non-IBL peers.

A similar pattern was found when survey data were differentiated by the number of prior college mathematics courses taken (3.2.5). Among IBL students, less experienced students reported higher gains than more experienced colleagues, but this was not the case for non-IBL students. Gains enhanced by IBL experience included cognitive, affective, and social gains for novice math students, but only affective and social gains for students with more math background.

In interviews, first-year math-track students reported more gains overall than advanced math-track and pre-service teaching students (Table 7.3; 7.4). Consistent with the survey results, they particularly emphasized cognitive gains, changes in their understanding of the nature of mathematics, and affective gains including confidence and enjoyment.

In sum, several lines of evidence indicate that IBL experiences were more powerful for students earlier in their college career. This finding is also consistent with data from an IBL first-year honors course on students' later course-taking patterns (Figure 6.9). A first-year IBL experience may contrast strikingly with students' high school work, and changes in students' approaches to learning or studying may influence their work in later courses in mathematics and other fields.

ES 4.4 Other Group Differences

When considering other student sub-groups, we found very few other differences, and no systematic patterns of difference. We detected no meaningful differences in the outcomes of IBL courses for students of varying race, ethnicity, or academic major. However, these study sites did not provide good tests of these issues. In general, we have no evidence that IBL methods did not work equally well for students of different personal and academic backgrounds.

ES 5.0 Student Attitudinal Outcomes

Student surveys explored several attitudinal variables that characterized students' mathematical beliefs, motivation and strategies for learning and problem-solving.

ES 5.1 Characterization of Students' Beliefs, Motivations, and Learning Strategies

Based on their pre-course survey responses, math-track students in both IBL and non-IBL classes had strong interest in mathematics, and high motivation in both intrinsic (internal) and extrinsic (grades, future plans) dimensions (3.4.2, Table 3.5). They held fairly sophisticated views of mathematical problem-solving as a constructive and logic-based process. While IBL and non-IBL students were alike in many ways, IBL students more often rated mathematics as a personal, not just academic interest. They were also more confident and had greater preference for group work. These differences confirm that there is some preferential selection (by advising and/or self-selection) of certain type of students into IBL courses, as students and faculty also told us.

Compared with math-track students, pre-service teachers were less interested in mathematics, less likely to enjoy it, and more extrinsically motivated (3.4.3). Their beliefs about learning and problem-solving were somewhat more novice-like, viewing learning as more instructor-driven and problem-solving as more about confirming truths and practicing procedures than about discovering ideas. However, they believed more in the value of group work, made more use of it in their own studying, and were more interested in teaching and communicating mathematics.

ES 5.2 Attitudinal Changes Following an IBL Course

In general, the changes in these attitudinal variables from pre- to post-course survey were modest for all groups (3.4.4). They were also modestly but positively correlated with student learning gains (3.4.13), showing that attitudinal changes and learning outcomes are indeed related.

For non-IBL math-track students, attitudinal changes were mixed, but more negative than positive (3.4.5). After their course, these students reported lower confidence and enjoyment, less willingness to study hard for a math course, and less strong beliefs in rigorous reasoning as a general problem-solving approach. But for IBL students, most of the changes were positive: stronger personal interest in mathematics and in communicating it, stronger beliefs in proving as a constructive and creative activity, and stronger beliefs in and use of collaborative learning.

Among pre-service teachers, attitudinal changes were mixed (3.4.6). Following an IBL course, they placed less emphasis on extrinsic goals and instructor-driven instruction, suggesting some maturation of their approach to learning mathematics. However, they did not gain in confidence, and lost some ground in their use of self-regulatory learning strategies that are used by successful learners.

For women, there were small positive changes in confidence and motivation following an IBL course, contrasting with larger negative changes in confidence, collaboration, and use of effective learning strategies for women who took traditionally taught courses (3.4.7, 3.4.8). These findings align well with the learning gains observed for women (ES 4.1): again, IBL approaches appeared to remedy problems with traditional lecture-based teaching that were particularly detrimental to college women's interest and confidence in mathematics.

Among first-year students, there were enhancements to students' strategies for learning and problem-solving, while these did not change for older students (3.4.9-3.4.11). But among older students, positive changes in interest, motivation and confidence were observed, which were instead modestly negative for the first-year students. We suggest that early IBL experiences have an influence on students' approach to learning that may be powerful if it carries over to other college work. Later IBL experiences may not shift students' well-established study habits and beliefs, but may revive their interest in mathematics, as some interviewees suggested (7.2.4).

Overall, IBL math courses tended to promote slightly more sophisticated and expert-like views of mathematics and more interactive approaches to learning. In contrast, traditional mathematics courses appeared to weaken students' confidence and enjoyment, and did not help them to develop expert-like views or skillful practices for studying college mathematics.

ES 6.0 Teaching and Learning Processes

Several lines of evidence clarify the teaching and learning processes important in IBL courses. Clear differences in IBL and non-IBL student outcomes (ES 3.0) mirrored the clear differences between IBL and non-IBL course practices (ES 2.2). More importantly, we can directly link student outcomes to course practices: student learning gains correlated statistically significantly with the fraction of class time spent doing student-centered activities (small group work, student presentation, computer work, and discussion), and anti-correlated with the fraction of time spent listening to instructors talk (4.3.1). Similar correlations were found for the relation of learning outcomes to the proportion of class time that was student- or instructor-led (4.3.1), and for variables that reflect how students and instructors interact and share responsibility for the course (4.4). Moreover, statistical modeling shows that the degree of student-centered class time was the strongest strong predictor of student learning as measured by our broadest learning indicator, survey learning gains. When observation data are not available, the binary IBL/non-IBL label was a good predictor of learning.

Students themselves reported in some detail on how particular course practices supported their learning. On surveys, IBL math-track students cited several practices as "helping me learn" to statistically higher degrees than cited by their non-IBL peers: the overall approach; their own active participation; and interactions with the instructor, the TA, and their peers (3.3.1, 3.3.2). Non-IBL students cited tests as important for their learning, while IBL students found other types of assignments more helpful. Individual effort was important to both groups' learning. IBL pre-service teachers emphasized a somewhat different mix of experiences, including their own active participation and interaction with instructors, but also tests.

Interactive and collaborative course experiences were especially important for women (3.3.3) and first-year students (3.3.4), which may help to explain the strong learning outcomes for these groups (9.4.1, 9.4.2). There were no clear patterns of difference in the experiences of other student sub-groups, consistent with the lack of clear patterns in their learning outcomes (9.4.4).

Finally, in interviews, student discussion of their learning processes emphasized the twin pillars of deep engagement with mathematical ideas and collaboration with others (7.3.1). Deep engagement fostered deep understanding; it rested on both students' individual effort and the assignment of meaningful problem-solving tasks that were not mere "busy work." Collaboration was integral to IBL courses, whether as structured small group work, whole-class discussion, or out-of-class informal group work. Students found it efficient and useful to tackle hard problems with multiple brains, and they learned from explaining their ideas and trying to understand others. The twin pillars reinforced each other: after struggling with a problem individually, students were well prepared to contribute meaningfully during class, and interested in the solutions that others proposed. Collaboration in turn motivated them to complete the individual work. It also made class enjoyable, encouraged clear thinking, and built communication skills.

Instructors also described the twin pillars and linked them to students' positive cognitive and affective outcomes (8.2.6). Descriptions of instructors' teaching practices enable us to identify critical teaching decisions that can sensitively influence the success of an IBL course. Their choices about course materials, assessment, classroom dynamics, and other factors may aid, abet, or interfere with the central learning processes of deep engagement and collaboration (8.3).

Overall, surveys and interviews provided strong and consistent evidence about the dual importance of individual engagement and collaborative learning processes in IBL courses. The significant role of collaborative learning reflects a deliberate shift in modern instructional practices that yielded enriched student learning, growth in collaboration and communication skills, and an enjoyable experience for students and instructors alike.

ES 7.0 Outcomes for IBL Instructors

Instructors reported numerous professional and personal benefits of teaching with IBL methods, which outnumbered their costs (8.4.1). The chief benefit was enhancements to their teaching: deeper understanding of students and learning; stronger beliefs in the value of student-centered learning; and a larger and more nuanced portfolio of teaching skills. They also gained intellectual stimulation, interest, enjoyment, and pride in their students' progress. The main costs were time and effort (8.4.2).

Early-career mathematicians felt that their career preparation and prospects were enhanced by IBL teaching experience (8.4.3). The more common career influence was professional development that they felt prepared them for future teaching roles. All who had gone on the job market reported that their IBL background had been an asset rather than a liability.

Of the instructors interviewed, at least 85% wanted to teach with IBL methods again (8.4.4). They described profound and permanent changes in their teaching styles and beliefs about what and how students learned. These influences on their teaching approach carried beyond the original courses where they had learned IBL methods, as they took "IBL principles" to other courses. Most of their learning came from on-the-job experience, although in most departments a loose-knit IBL community had emerged that could offer support and ideas.

Overall, IBL teaching experiences were rewarding for instructors. As professional development experiences, they permanently shifted instructors' beliefs and practices toward student-centered approaches known to improve student learning. The most important legacy of the project may be this cadre of young instructors who gained IBL teaching experience at the Centers and are now moving on to teach at institutions across the United States and in other countries.

ES 8.0 Outcomes for the Project as a Reform Effort

Outcomes for students and teachers are one way to view the overall impact of the IBL Centers. We can also view the Centers' work collectively and examine the magnitude of its impact as a single reform effort: How big an impact does this project have?

About 425 unique, non-repeating students had an IBL experience each year during our study period. Over the same period, these four departments graduated about 500 mathematics majors each year.[4] On the order of 40% of all mathematics majors graduating from the four Centers may have had an IBL experience. The two institutions with IBL programs that targeted pre-service teachers graduated about 160 pre-service teachers per year.[4] Essentially all students preparing for elementary/middle school teaching at these two campuses had an IBL experience.

Another measure of impact is the sustainability of the IBL Centers' effort over time. All the Centers had taken measures to protect aspects of their program from budget cuts. Some aspects of the IBL programs were seen to be "here to stay," yet leaders also testified to aspects that were fragile and dependent on grant funding. In times of economic retrenchment and falling support for U.S. higher education, the jury is still out as to the sustainability of these reform efforts.

ES 9.0 Issues for Future Research

Our literature reviews identify fairly modest bodies of evidence on student outcomes of IBL for mathematics majors and for pre-service teachers. Findings from this study will thus contribute meaningfully to these literatures, as well as add to important conversations on teaching practices in college mathematics and on change in STEM higher education. We identify several issues for further analysis within our data sets and propose several areas for future investigation (9.9)

ES 10.0 Conclusion

The approaches implemented at the IBL Mathematics Centers benefited students in multiple, profound, and perhaps lasting ways. Learning gains and attitudinal changes were especially positive for groups that are often under-served by traditional lecture-based approaches, including women and lower-achieving students. First-year and less mathematically experienced students also benefited particularly. Yet there was no evidence of negative consequences of IBL for men, high-achieving students, older and more experienced students: these groups too made gains greater than their non-IBL peers.

---

[4] National Center for Education Statistics (NCES) (2011). Integrated Postsecondary Education Data System (IPEDS). Department of Education. Office of Educational Research and Improvement.

The positive outcomes for students were linked to classroom practices that emphasized deep engagement with mathematical ideas and collaborative exploration of these ideas. IBL classrooms offered equal learning opportunities for men and women and motivated students to invest their own effort to advance class progress. Instructors also benefited from their IBL teaching experiences and made lasting changes to their teaching practice. On the whole, the teaching and learning methods implemented at the IBL Mathematics Centers were broadly consistent with evidence and best practices from research on the learning sciences. Our results augment that body of evidence with good support for student-centered approaches to undergraduate mathematics education.

ES 11.0 Acknowledgments

Many individuals provided assistance that made these findings possible (1.7). We thank the Educational Advancement Foundation (Austin, TX) for support of this work. All conclusions are the authors' own and do not necessarily reflect the views of the EAF.

Table of Contents

Appendices Online at http://www.colorado.edu/eer/research/steminquiry.html#Reports

List of Tables

List of Figures

Chapter 1: Introduction to the IBL Mathematics Project, the IBL Assessment & Evaluation Study, and this Report

1.1 Introduction

We report findings from a large, mixed-methods study of inquiry-based learning (IBL) in college mathematics. The study is comprised of six linked sub-studies of inquiry-based and comparative courses that were developed and taught at four university IBL Mathematics Centers. The study was designed to examine the following questions:

1) What are the student outcomes—including learning, attitudes, beliefs, career and education plans—of IBL mathematics courses?
2) How do these outcomes vary among student groups, and how do they compare with other types of courses?
3) How do these outcomes come about? What roles are played by students, instructors and teaching assistants, course materials, assignments, assessments, and other practices?
4) What are the costs and benefits for instructors and departments who teach with IBL methods?

We begin this introductory chapter by describing the purpose, audience and organization of the report. We then locate the study in several important contexts: a policy context of national concerns about mathematics education, a local context within the work of the IBL Centers and the Educational Advancement Foundation (EAF), and an intellectual context of prior research on inquiry-based learning in undergraduate mathematics. We sketch the design of the study, both as a whole and in its six components. Last, we define some terms used throughout the report.

In Chapters 2-8, we report selected results from our analyses, highlighting findings on student outcomes in IBL and non-IBL courses. Each chapter focuses on results from a particular sub-study. Chapter 9 summarizes findings grouped by theme across multiple lines of evidence.

1.2 Purpose and Organization of this Report

In this report, we focus strongly on student outcomes—that is, on Research Questions #1 and #2—and we address the process-focused Questions #3 and #4 only briefly. Our data sets are quite extensive, and this report highlights selected analyses that are both complete at this time and relevant to the current work of the EAF and IBL Centers.

1.2.1 Audience

This report is written as an internal evaluation report to the funders and participants in the project. As such, it is directed to a practice-focused audience, which may additionally include:

- Mathematics instructors and departments considering whether and how to incorporate IBL teaching and learning methods into their programs
- Students considering whether to take an IBL course

- Leaders situated in universities, foundations, and policy-making bodies who are considering whether and how to support and disseminate IBL methods.

To meet their needs, we focus on the findings and the supporting evidence and limit the use of technical vocabulary. We highlight key findings and point to others—all in a context of a very large volume of data gathered and findings to share. We describe the research methods very briefly and refer readers to the appendices for methodological details. We invite readers to ask questions and to tell us where we have or have not succeeded in meeting their needs.

At the same time, however, we believe that it is important for non-technical readers to learn not just "what we know" but to gain some understanding of "how we know what we know." Otherwise, our claims will sound no different than anecdotes, nor will they add anything new to the conversation about IBL. Indeed, some of our findings may seem little different in character from what experienced IBL teachers have observed for themselves and shared with colleagues—but the research-derived claims we present are importantly different in their origins. Our evidence is systematically gathered from a large number of participants with differing perspectives, and with attention to bias and self-selection. Such data are not inherently better or more useful than personal experience, but they do have a different weight and level of generalizability than do individual accounts or observations.

1.2.2 Limitations to the Scope of this Report

In addition to the primary, practitioner audience, the study findings will also interest researchers and evaluators of undergraduate education and teacher preparation in science, technology, engineering and mathematics (STEM) fields, and scholars of educational reform. For this audience, a set of substantial appendices on the study samples and methods will, we hope, provide the information needed to assess the study's methodological rigor.

Second, for brevity, we do not compare our findings with the literature. We have reviewed several bodies of relevant literature and summarize them in Section 1.4, but we do not relate individual findings to what is already known. We will address the linkages between our and others' work in future publications.

Finally, the results presented here are a subset of first findings. We designed these studies to provide multiple, complementary measurements, but have not yet exploited the potential connections among them. Analyses cross-linking multiple lines of evidence will be essential to clarify complex phenomena such as learning and teaching processes in IBL classrooms, their relationship to student outcomes, and the experiences of important groups, such as pre-service teachers. Future publications will explore these issues in greater detail. We have many ideas for future analyses and welcome questions or ideas from others.

1.2.3 Organization of the Report

The report is organized simply: results from the six sub-studies are provided separately, each in their own chapter. One extra chapter (4) addresses the linkages between two sub-studies, the student surveys and classroom observations. This narrative strategy enables us to share certain findings rapidly, despite the massive scale of the data sets and the many analyses yet to be completed. Chapter 9 summarizes findings across sub-studies. We encourage readers to consult the Table of Contents to find topics of interest to them.

A set of detailed appendices carries details about the study samples and methods; they are referenced throughout the report. We intend that these details lend credibility and transparency to the findings without interrupting the narrative. Some tools that we have developed may be useful to evaluators or researchers, such as protocols and survey instruments. Because they are lengthy, the appendices are not part of this document, but can be obtained at our web site, http://www.colorado.edu/eer/research/steminquiry.html

1.3 National and University Context for the Study

Interest in inquiry-based learning is driven in part by concerns about the quality of American mathematics education. Policy-makers seek to ensure a strong U.S. workforce in science, technology, engineering, and mathematics, collectively known as the STEM fields. For individuals, a good mathematics education offers the possibility to pursue good jobs in these fields. We describe these issues broadly, then describe the response of the EAF and the IBL Centers to this need, which provided the research context for this study.

1.3.1 National Concerns about Mathematics Education

The concern to develop American citizens who possess strong mathematical reasoning abilities and critical habits of mind is not new (Colburn, 1826; Dewey, 1938; NCTM, 1989, 1991; NRC, 1989; Paulos, 1988, 1995; Leitzel, 1991). But in recent years this issue has escalated as a national policy concern. The quality and quantity of the STEM workforce crucially affects national economic security, global competitiveness, and future innovation (COSEPUP, 2007). Since 1950, the number of nonacademic science and engineering jobs has grown at nearly four times the rate of the U.S. labor force as a whole (NSB, 2010), and these jobs are argued to have high leverage throughout the economy by spurring innovation and job growth (Members of the 2005 "Rising Above the Gathering Storm" Committee, 2010). By training mathematicians and supporting other STEM fields, strong K-20 mathematical preparation helps to maintain this sector of the economy (COSEPUP, 2007).

Moreover, increasing the diversity of the STEM-trained workforce has been identified as “an urgent need,” given changing demographics, decreasing numbers of foreign citizens entering the U.S. STEM workforce, and growing international competition for scientific and engineering talent (CEOSE, 2005). Concerns about equity and justice demand that all Americans have equal opportunities to enter science and technical careers, which are typically high in status and pay and less often subject to unemployment (National Science Board (NSB), 2010; Chapter 3 Highlights). Yet such equity has not been achieved: In mathematics itself, women’s share of bachelor’s degrees has declined in the past decade and a half (NSB, 2010; Chapter 2, Figure 2-

6), while the share of bachelor's degrees awarded to underrepresented minority groups[1] has remained nearly flat (NSB, 2010; Chapter 2, Appendix Table 2-13). Providing all Americans with access to high-quality mathematics education is critical in meeting workforce needs and providing equal opportunity for all citizens.

Inquiry-based learning (IBL) is argued to help address these issues. Rather than emphasizing rote memorization and computation skills, IBL approaches seek to help students develop critical thought processes: analyzing ill-defined problems, weighing evidence, applying logic, making and analyzing arguments. The same mental activities also support deep learning of mathematical ideas. And, by building students' confidence in their abilities to generate and critique ideas and to solve problems independently, IBL methods may foster students' creativity and persistence. EAF and the IBL Centers seek to address these concerns at multiple levels, hoping to:

- foster the success of undergraduates who will join the STEM workforce;
- develop research talent among students who will go on to become STEM innovators;
- close the gap between populations traditionally successful in mathematics and those who have been historically underrepresented;
- enhance critical thinking and intellectual growth among college graduates who will become voters and leaders in business, civic and community life; and
- improve the mathematical training of K-12 teachers who will educate the next generation.

It is these broad aims that frame the practical and policy significance of our study findings.

1.3.2 Local Context for the Study: The Work of the IBL Mathematics Centers and EAF

Our study was commissioned by the Educational Advancement Foundation of Austin, Texas, to investigate student outcomes from courses developed by the IBL Mathematics Centers also funded by EAF. Four of these university IBL Centers were established in 2004 in the mathematics departments at the University of Chicago, University of Michigan, University of Texas at Austin, and University of California, Santa Barbara (UCSB). A fifth IBL project, more limited in scope, began at Harvard University at the same time. The purpose of the Centers was "to further develop, study, promote and disseminate the use of inquiry-based learning (IBL) approaches in the teaching of mathematics by fostering IBL activities at five prestigious national universities" (EAF, 2007).

As the quotation shows, it was thought that these leading research departments would have high visibility and influence in their discipline. Each also had some history of engagement in mathematics education. For instance, Harvard and Michigan were leaders in the calculus reform effort of the 1990s. Chicago and UCSB had been active in K-12 standards, curriculum, and teacher professional development in their home states. And Texas was the home of mathematics

[1] In mathematics, as in most STEM disciplines, underrepresented racial and ethnic groups include African-American or Black, Hispanic or Latino/a, and Native American/Native Alaskan students.

professor R. L. Moore (1882-1974), renowned teacher and topologist whose socratic "Moore method" inspired EAF's work and remained (in modified form) in active use.

Working in parallel, the Centers have developed a variety of undergraduate mathematics courses that emphasize some common pedagogical elements. Students solve difficult problems on their own or in groups; they present their solutions at the board; and they critique and refine each others' solutions in class discussion. The mix of activities may emphasize student presentations or structured group work, depending on the student audience and the instructor's choice. But in each case, students' ideas and explanations provide the core material of the course and drive progress through the curriculum. Largely independent of a textbook, course content is instead based on a carefully built sequence of problems that lead students in small steps to the big mathematical ideas of the course. Instructors' learning goals for their students thus tend to place equal emphasis on building deep understanding of mathematical ideas, and on developing critical habits of mind useful in both mathematics and life in general.

A visitor to an IBL class would see students not passively taking lecture notes, but working in groups or presenting problems at the board, guided by faculty and TAs and critiqued by their peers. Compared to a lecture course, instructors play a crucial but less visible guiding role in setting the sequence of problems and shaping the discussion at key moments. While "inquiry-based learning" may sound somewhat generic, the IBL Centers share a rich heritage that includes the socratic teaching practices of R. L. Moore (see also Jacob, n.d.). Today Moore's approach is importantly modified to incorporate greater interaction among students, foster a positive learning environment in the classroom, and welcome a broader talent pool.

Within this shared conceptual framework and pedagogical heritage, each department developed its own menu of IBL mathematics courses for undergraduates. The courses fit the opportunities, constraints, and staffing available on each campus, and collectively illustrate a wide array of IBL applications.[2] At Chicago, a three-quarter sequence in first-year Honors Calculus, and a three-quarter sequence in sophomore Analysis have been mainstays, while IBL courses in Number Theory and Geometry are more recently developed. At Michigan, core offerings have included Introduction to Cryptology for first-year students; Analysis for upper-division students; and three courses for pre-service K-12 teachers: a two-semester sequence required of K-8 certification students, and a one-semester elective for pre-service middle school teachers.

At Texas, core courses have included Number Theory, two Analysis courses, and occasional sections of Discrete Math and Geometry, all at the sophomore to senior levels. At UCSB, core courses include a three-quarter honors sequence in Multivariable Calculus, Linear Algebra and Differential Equations, for first- and second-year students. Two two-quarter sequences separately target elementary and secondary pre-service K-12 teachers. Harvard has experimented with undergraduate courses at the introductory and middle level, such as Discrete Math and

[2] Campuses also supported a variety of professional development activities, such as workshops for in-service K-12 teachers, TA training, and brownbag lunch gatherings for faculty. Only the undergraduate courses were direct subjects of this study, although we noted interactions with the other activities.

Introduction to Proof. While not exhaustive, these lists highlight the wide range of courses offered and the variety in student audiences addressed. Other than some sharing of course syllabi and "scripts," or sequences of problems or theorems to be proved, the courses were developed quite independently.

In addition to its direct support of the IBL Centers, EAF sponsors several other activities to promote IBL approaches generally as well as support and publicize the Centers' work. EAF held the first *Legacy of R. L. Moore* conference in 1998. Today this meeting draws over 300 participants annually, mostly college mathematics educators interested in the "modified Moore method" and other forms of inquiry-based learning. EAF's mentoring and small grants programs have supported dozens of projects across the country. At least 200 faculty (including a few IBL Center instructors) have participated in intensive, multi-day workshops sponsored by EAF or its collaborators. Connections have been forged with other faculty development efforts such as the Mathematical Association of America's Project NExT (New Experiences in Teaching) and Professional Enhancement Program (PREP). EAF also distributes a body of recorded and written materials, such as historical and personal accounts of inquiry-based learning. The Academy for Inquiry-Based Learning provides an infrastructure for linking practitioners to each other and to all these resources. Together, these efforts have built a community of practice that connects scattered individuals and welcomes new members into the fold.

After the IBL Centers were established, EAF recognized a need for assessment and evaluation data that would offer evidence about the effectiveness of IBL methods (EAF, 2007). Our group, Ethnography & Evaluation Research, proposed and carried out a two-phase study that was funded by EAF and given a formal designation as the IBL Mathematics Assessment and Evaluation Center. First, we conducted a nine-month planning study, touring the five campus sites to assess needs and opportunities and to gather input from students, teaching assistants (TAs), faculty and campus leaders on issues and outcomes of interest to them. Based on this initial study we designed a comprehensive study to examine student outcomes and teaching and learning processes in these courses.

Four campuses elected to participate in the comprehensive study,[3] which was funded by the EAF and carried out over two academic years, 2008-2010. These data are the source of the findings presented here. We hope that the results will not only document the student and instructor experience of these methods and suggest ways to measure their effectiveness, but also help IBL instructors to refine their methods, overcome barriers, and inform their students and colleagues.

1.4 Intellectual Context for the Study: Overview of the Literature

In this section, we offer two forms of intellectual grounding for this study. First we discuss the theoretical and conceptual basis for IBL methods as applied to undergraduate mathematics. Second, we summarize prior empirical studies of inquiry-based learning and related approaches in undergraduate courses taught for mathematics majors and, separately, for pre-service teachers.

[3] Harvard chose not to participate in the study, and we do not discuss the Harvard IBL project further.

1.4.1 Theoretical and Conceptual Basis for IBL Methods

While IBL methods arose from practitioner experience, through the spread and adaptation of Moore's teaching methods by his former students, they are consistent with the current scholarly understanding of how people learn. Here we place IBL in the context of the broader theoretical and empirical research base behind inquiry learning.

Inquiry-based learning refers to the "scientific method," the process by which researchers create new knowledge. In giving students the chance to engage in knowledge creation, IBL promotes independent thinking and intellectual growth (e.g., Buch & Wolff, 2000). Students think and act like mathematicians as they work with ideas and formulate solutions to open-ended problems.

Like "discovery learning" (Bruner, 1961; Dewey, 1938), "problem-based learning" (e.g., Savin-Baden & Major, 2004), and other "inductive teaching" approaches (Prince & Felder, 2007), IBL invites students to work out ill-structured but meaningful challenges. By drawing on their own experience and prior knowledge, exploring their environment or performing experiments, and wrestling with questions and controversies, students learn what they need to know to address the challenge. In this process, students build critical thinking, analysis, and communication skills and learn to use resources efficiently (e.g., Duch, Groh & Allen, 2001). These methods vary in the form and extent of support and structure provided to students.

IBL is also a form of "active learning," the goal of which is to engage students in the learning process, and thereby activate responsibility for their own learning processes (Prince, 2004). Students' intention to learn, direct experiences of learning, and reflection on those experiences combine to yield deeper understanding (e.g., Moon, 2004; McCann, Johannessen, Kahn & Smagorinsky, 2005). Such learning becomes "transformational" when it shapes people so profoundly as to affect all their subsequent learning (Mezirow, 1991).

As implemented in this project, inquiry-based learning draws on all these features. In particular, inquiry methods encourage students to take responsibility and enhance higher-order thinking skills. As students deepen their mathematical knowledge, they also become more effective thinkers and learners. And gains in metacognitive or self-regulatory skills in turn foster crucial, general academic skills such as problem-solving and creativity (e.g., Zimmerman & Schunk, 2001). Self-regulated learners approach tasks strategically, with diligence, confidence and resourcefulness—and inquiry can enhance these capacities.

While Moore emphasized individual learning, modern implementations of IBL in mathematics draw on the so-called social learning perspectives (e.g., Lave & Wenger, 1991; Vygotsky, 1978) that are importantly applied in mathematics education (e.g., Alrø & Skovsmose, 2002; Cobb, Yackel & McCain, 2000; Lerman, 2000). In this view, knowledge is shared and constructed through social interactions. Teaching approaches that emphasize social learning most often draw on the ideas of cooperative or collaborative learning. While they vary in how group work is structured and evaluated, and in the authenticity of the group task, in general such approaches foster student dialogue, build positive interdependence among group members, challenge students' thinking so as to promote higher-order thinking, and encourage students to accept responsibility for their own learning (e.g., Goodsell, Mahler, Tinto, Smith & MacGregor, 1992; Johnson, Johnson & Smith, 1998). These approaches thus yield high educational gains and improved social skills (Gillies, 2007; King, 2002).

Overall, inquiry learning is supported by a strong theoretical rationale that draws on several important elements of active learning. Inquiry and cooperation create an efficient and enjoyable learning context for mathematics students. IBL approaches thus lay a path to mathematical empowerment (Manigault, 1997) that extends well beyond college classrooms. As Perkins and Tishman (2001) suggest, these approaches develop habits of mind that provide young people with good leverage on modern challenges: a disposition toward wondering, problem-finding and investigating, and the ability to be metacognitive and intellectually careful.

1.4.2 Prior Studies of Student Outcomes of IBL in College Mathematics

To understand how our results relate to the landscape of previous work, we commissioned two literature reviews. The first focused on IBL methods as applied to college mathematics at the calculus level or higher (Hough, 2010a). The search began with the Modified Moore Method most closely related to the IBL Centers' work, then broadened to incorporate literature on other inquiry-oriented and active learning approaches. While lower-level courses such as college algebra, statistics, and general education courses offer important, high-impact venues for potential application of IBL approaches, none of the IBL Centers developed such courses, and we excluded them from our literature review. Our review focused on student outcomes, but also attended to outcomes for instructors and departments and analyses of learning processes.

The results of this search indicate that to date, very little evidence exists past the calculus level of the effects of inquiry-based learning on students, instructors or the departments in which IBL is implemented. Several mathematicians have written thoughtful descriptions of the Moore Method (or Modified Moore Method, MMM) as implemented in their college classrooms (e.g., Buck, 2006; Dumitraşcu, 2009; McLoughlin, 2009). But we found only four empirical studies of the Modified Moore Method in college mathematics, and only two were methodologically rigorous.

For this search, Hough (2010a) used a rigorous approach, searching in the ERIC, JSTOR, Sage Online Journals and ProQuest databases using all three-word combinations of the following search terms: *mathematics, STEM, college, higher education, undergraduate, Moore, Method, Modified Moore Method, inquiry, inquiry-based learning, guided re-invention, discovery learning, small group learning, cooperative learning, collaborative learning, problem-based learning, project-based learning*. She also reviewed all abstracts back to 2000 in nine journals that publish mathematics education research. This procedure uncovered a relatively small body of evidence, which was classified into four strands as shown in Table 1.1.

While not all use the term inquiry-based learning, the studies included here examine courses that clearly share many features of the practices used by the IBL Centers. Courses based on the philosophy of "realistic mathematics education" (RME) (Freudenthal, 1991; Gravemeijer, 1994) similarly emphasize students' active creation of and responsibility for the mathematical concepts. Class moves forward as students reach understanding, rather than being pegged to a pre-set schedule, and often include small group work and student presentations.

Table 1.1: Studies of Inquiry-Based Learning in College Mathematics for Math-Track Students

Category of study	Number of studies (comparison group)	Study Topic: student achievement, understanding, or skills	Study Topic: student attitudes & beliefs	Study Topic: long-term effects (e.g., persistence & grades)
Modified Moore method	2 (C)	Christian Smith, 2006 Jensen, 2006		Jensen, 2006
realistic mathematics education	4 (no C)	Rasmussen & Blumenfield, 2007 Larsen, 2008 Larsen & Zandieh, 2009	Ju & Kwon, 2007	
	2 (C)	Kwon, Rasmussen & Allen, 2005	Kwon, Rasmussen & Allen, 2005	Rasmussen et al., 2006
calculus reform	11 (C)	Bookman & Friedman, 1994, 1999 Schwingendorf, McCabe & Kuhn, 2000 Penn, 1994 Williams, 1998 Meel, 1998 Roddick, 1993 Ganter & Jirovtek, 2000 Hurley, Koehn & Ganter, 1999 Smith & Star, 2007	Bookman & Friedman, 1994, 1999 Smith & Star, 2007	Bookman & Friedman, 1994, 1999 Schwingendorf, McCabe & Kuhn, 2000 McDonald, Mathews, & Strobel, 2000 Hurley, Koehn & Ganter, 1999
specific instructional activities typical of IBL mathematics courses	6 on small-group learning (C)	Treisman, 1986 Bonsangue, 1994 Bonsangue, 1991 Herzig & Kung, 2003 Urion & Neil, 1992 Springer, Stanne, & Donovan, 1999	Springer, Stanne, & Donovan, 1999	Treisman, 1986 Bonsangue, 1994 Herzig & Kung, 2003 Springer, Stanne, & Donovan, 1999
	1 on student presentations (C)	Alsardary & Blumberg, 2009		

C: results were compared with a non-IBL comparison group. no C: no comparison group was studied.

The U.S. calculus reform movement of the 1980s-90s (Tucker, 1995) emphasized, to varying degrees, conceptual over procedural knowledge of mathematical ideas and sought to actively engage students in constructing those concepts using authentic processes. Reform classrooms often incorporated collaborative group work and technology and broadened student tasks to include real-world scenarios, writing and reflection. Lastly, a few studies have examined specific instructional methods common in IBL courses. Springer, Stanne, and Donovan's (1999) meta-analysis of small-group work in college STEM courses is a classic of the field, but no studies of small-group work in undergraduate mathematics met their criteria for inclusion.

Based on these four strands of studies, we observe several patterns in student outcomes. Among studies of achievement and understanding, it is clear that the cooperative learning workshop model supports the success of underrepresented students in traditional calculus courses (Treisman,1986; Bonsangue, 1994). But simply implementing cooperative learning techniques or adding a computer laboratory to a traditional setting without changing how students interact with the mathematics does *not* appear to change student outcomes (Ganter & Jirovtek, 2000; Herzig & Kung, 2003; Urion & Neil, 1992). When students take courses that focus on authentic mathematical inquiry, they are able to construct for themselves important mathematical ideas (Christian Smith, 2006; Larsen, 2008; Larsen & Zandieh, 2009; McDonald et al., 2000). Such reconstruction of mathematical ideas allows students in IBL courses to develop deeper conceptual understanding (Christian Smith, 2006; Kwon, Rasmussen & Allen, 2005; Rasmussen, et al. 2006; Williams, 2008) and to solve problems more efficiently (Bookman & Friedman, 1994, 1999; Meel, 1998) than their peers in traditional courses. Students who used problem-based learning score at least as well and often higher than students who take traditional courses, as measured by grades (Schwingendorf, McCabe & Kuhn, 2000), final exams (Penn, 1994), and assessments of procedural skills (Meel, 1998).

There is less evidence that IBL approaches change student attitudes toward mathematics in general (Bookman & Friedman, 1994, 1999; Smith & Star, 2007). But there is some evidence suggesting that female students benefit more from participating in active learning or guided reinvention environments than do male students (Hurley, Koehn & Ganter, 1999).

In general, studies that look at the long-term effects of IBL in terms of grades, course-taking patterns, persistence or conceptual understanding have showed no differences compared to students taking traditional courses, or have given mixed results. Studies of IBL students' problem-solving skills yield more promising results: IBL students retain advantages in solving open-ended calculus problems after one year (Bookman & Friedman 1994, 1999) and have greater ability to solve engineering problems involving calculus in their "next" course.

1.4.3 Prior Studies of Outcomes of IBL for Pre-Service Teachers

Inquiry-based approaches are seen as particularly important for pre-service teachers, because the ability to understand why and how certain mathematical rules work is a critical component of the knowledge needed to teach mathematics (Conference Board of Mathematical Sciences, 2001). According to the Conference Board, pre-service education should rekindle prospective elementary teachers' own powers of mathematical thinking "with classroom experiences in which *their* ideas for solving problems are elicited and taken seriously, their sound reasoning affirmed, and their missteps challenged in ways that help them make sense of their errors" (p. 17). And secondary teachers need opportunities to make insightful connections between the advanced topics that they are learning and the high school topics that they will be teaching.

In addition to subject matter, mathematical knowledge for teaching (MKT) also includes deep understanding of how ideas and concepts connect to one another, and of how children develop understanding of these ideas: their prior knowledge or misconceptions, the representations and strategies that children create and teachers can build upon to foster understanding. Building on Shulman's (1986) notion of pedagogical content knowledge, MKT is that special amalgamation of content and pedagogical knowledge a good teacher draws on to make instructional decisions. MKT is one of three key areas identified as integral to improving student outcomes (RAND, 2003). Ball and colleagues distinguish MKT from the mathematical knowledge held by other professionals (Ball, Hill & Bass, 2005) and have devised a measure of it, the *Learning Mathematics for Teaching* (LMT) instrument, that has been used to show a positive relationship between teachers' MKT and their own students' achievement in mathematics (Hill, Ball & Schilling, 2008; Hill & Ball, 2004). To date, we have found no published studies that use the LMT in order to test whether and how this type of knowledge develops in prospective teachers, although some studies explore this using other measures.

Using the same databases and keywords, plus the additional terms *prospective teacher* or *pre-service teacher*, and searching the same journals, Hough (2010b) scanned the literature on learning outcomes for pre-service teachers taking inquiry-oriented courses. We included both undergraduate mathematics courses specifically designed for prospective teachers and teaching methods courses that emphasize the development of mathematical ideas. Surprisingly, while there is a broad research base on prospective teachers' difficulties with mathematics, very few studies evaluate courses or programs to remedy this problem. Moreover, of the nineteen outcomes-oriented studies identified by Hough, only three offer data comparing outcomes from inquiry-based or active learning courses to those from more traditional formats.[4] These studies are listed in Table 1.2.

Studies of prospective teachers' <u>subject matter knowledge</u> generally find that IBL approaches help prospective teachers to make gains in specific subject matter knowledge (e.g., proportional reasoning, arithmetic operations). The single comparative study found that students who took a problem-based course outperformed the traditionally taught group on a researcher-developed test of Cantorian set theory (Narli & Baser, 2008).

Among studies of students' understandings of the <u>nature and processes of mathematics</u>, authors report generally positive outcomes. For example, students learn to participate in mathematical discourse and develop more sophisticated conceptions of creativity in mathematics. In the only comparative study, Yoo and Christian Smith (2007) surveyed students who took Modified Moore Method or lecture-based courses on proof. The MMM students developed a significantly more humanistic and process-oriented view of proof than did their peers in the standard course.

---

[4] Because some studies measured outcomes fitting more than one category, the total number of studies is less than the sum of the three categories.

Several studies offer data related to changes in prospective teachers' beliefs, attitudes and efficacy surrounding mathematics. Authors report some gains in confidence and modest shifts in students' beliefs about how mathematics should be taught and learned. In a single comparative study, students who took an active-learning algebra course focused on learning and analyzing alternative algorithms held more positive attitudes about mathematics and (when controlled for GPA) had developed greater MKT as measured by a researcher-developed scale (Mathews & Seaman, 2007).

Table 1.2: Studies of Inquiry-Based Learning in College Mathematics for Pre-Service K-12 Teachers

Category of study	**Number of studies (comparison group)**	**Study Topic**			
		subject matter knowledge, including MKT	**understanding the nature & processes of math**	**beliefs, attitudes, or confidence about math**	**classroom practice in later teaching position**
single college courses for pre-service teachers	16 (no C)	McClain, 2003 Ben-Chaim, Keret, & Ilany, 2007 Chapman, 2007 Lo & Grant, 2007 Hough et al., 2011	Banton, 2002 Stylianides & Stylianides, 2009 Bolden, Harries & Newton, 2010 Shriki, 2010	Enemaker, 1995 Owens et al., 1998 Hart, 2002 Szydlik, Szydlik & Benson, 2003 Wilkins & Brand, 2004 Ben-Chaim, Keret & Ilany, 2007 Lubinski & Otto, 2010	
	3 (C)	Narli & Baser, 2008 Mathews & Seaman, 2007	Yoo & Christian Smith, 2007	Mathews & Seaman, 2007	
large-scale reform efforts in teacher preparation	3 (C)			McGinnis et al., 2002	Lawrenz, Huffman & Gravely, 2007 Adamson et al., 2003

Three studies report on evaluations of large-scale initiatives to improve undergraduate teacher education, the NSF-funded Collaboratives for Excellence in Teacher Preparation (CETPs). Two of the studies followed undergraduates into the classroom after graduation and report positive differences in the classroom practices of new, CETP-trained teachers (Lawrenz, Huffman & Gravely, 2007; Adamson et al., 2003). The third study reported positive changes over time in CETP teacher candidates' "reform-oriented" attitudes—but negative impacts on the attitudes of program non-participants who took the same courses (McGinnis et al., 2002).

All of the findings suggest that active learning experiences have a positive effect on teachers' content learning, understanding of the nature and processes of mathematics, and attitudes. Content knowledge may be the easiest of these to affect, and attitudes the trickiest. However, it is difficult to draw firm conclusions given the small size of most study samples and the dearth of carefully controlled comparative studies.

1.5 Study Design

The overall project addresses the broad question, "Does IBL work? If so, under what circumstances? If not, why not?" We sought to examine a range of student learning and affective outcomes—mathematical content knowledge, problem-solving skills, higher-order thinking, independent and collaborative learning, confidence, persistence, and interest—as well as longer-term impacts on students' education and career paths. To understand how such outcomes might come about, we studied the classroom context and the teaching and learning processes that took place in and out of class. We wished to understand the plusses and minuses for instructors of teaching this way, and the departmental and institutional factors that influenced their choices and the success and sustainability of the IBL programs on each campus.

Toward these aims, the study was designed as a checkerboard of sub-studies, shown in Figure 1.1. Not every research method could be used in every class, but together they build a detailed picture of where and how IBL methods do and do not "work" for students and their instructors.

Figure 1.1: Design Matrix for Investigation of Student Outcomes:
Approaches for Examining Outcomes within Key Student Audiences

	Math learning & thinking: Tests	Math learning & thinking: Self-report	Attitudes & beliefs	Career & educational outcomes	Classroom processes
Math, science, & engineering majors	Proof test*	Survey* Interviews	Survey* Interviews	Academic records* Interviews	Observation* Interviews
Pre-service K-12 teachers	LMT test	Survey Interviews	Survey	Interviews	Observation Interviews

* Comparative data was gathered from non-IBL sections of some courses.

Classroom observations provide a foundation for the other studies, enabling us to describe the teaching methods in use and link student outcomes to particular teaching approaches or classroom traits. Surveys, tests, transcript analysis, and interviews allow us to probe both student outcomes and learning processes with multiple methods. "Post" (end of course) measures focus on student outcomes, while "pre" (start of course) measures let us assess whether students are selectively choosing (or being advised) into and out of IBL courses. Interviews with faculty and TAs capture their observations of student outcomes and crucial perspectives on instructors' goals and methods, and document the costs, benefits and career impact of teaching an IBL course.

Most of the Centers' IBL courses targeted what we call "math-track" students, those enrolled in challenging mathematics courses. Most of these were mathematics, science and engineering majors. Others were first-year students confident in their math abilities but not yet committed to a major, taking honors-style courses targeted to them. Using surveys and interviews, we could explore these students' classroom experiences and their learning and attitudinal gains (or lack of gain) from their courses, and probe their intentions for future study or work in mathematics. Using student academic records, we looked for patterns in their grades and course choices after taking an IBL course. Given the great variety of courses and audiences, external tests were not readily applied, but we did gather and analyze some student work. For some courses, data from non-IBL sections provided a comparison group.

Two Centers applied IBL methods to several courses designed specifically for pre-service K-12 teachers, who are an important student audience quite distinct from math-track students. As well as exploring their learning goals and classroom experiences, surveys and interviews enabled us to examine these students' attitudes and beliefs about studying mathematics. Research has shown that teachers' attitudes and beliefs crucially affect their own students' later success (e.g., Beilock, Gunderson, Ramirez & Levine, 2010). To measure these students' content competency for teaching, we used a nationally validated pre/post test that focuses on mathematical concepts taught in K-12 grades and encountered in classroom situations. No comparative non-IBL sections were available for these IBL courses targeting pre-service teachers.

To meet our ethical obligations as researchers and regulatory requirements for human subjects research, all studies were submitted to and approved by the Institutional Review Board at the University of Colorado at Boulder and/or the study campuses. For example, interview, test and survey subjects provided informed consent before participating in these studies.

Next we sketch each of the sub-studies individually. Our aim here is to provide the general background required to understand the study results. Readers seeking additional methodological details should consult the Appendices at http://www.colorado.edu/eer/research/steminquiry.html

1.5.1 Classroom Observation

Classroom observation is a linchpin of the overall study because it lets us characterize and compare the teaching and learning approaches used across classrooms. Only with these data can we try to explain the effectiveness of different classroom approaches and variations within them. In comparing IBL and non-IBL courses, we followed each campus' own designation of their courses, but observation enabled us to ground-truth the extent to which these courses were truly student-centered or "inquiry-based" when so designated. Thus this study is critical for interpreting all the others.

Because we could not be physically present at all the campuses and classes at once, we hired and trained local observers who had expertise in advanced mathematics education. They attended six to ten sessions of each course over the term and used a simple rubric to record how class time was spent and how students and instructors interacted during class. Observers categorized the type of classroom activity taking place and the number of minutes it lasted. They tallied the number of questions asked and recorded the identity of the question-asker (student, faculty, TA) and the nature of the question (e.g., simple recall or higher-level analysis). They completed a survey about interpersonal interactions and atmosphere in the classroom and took notes to describe what they saw. The observation rubric focused on the approaches used to engage students in course material, rather than on instructors' skill in implementing these methods.

To develop the rubric, we drew upon our classroom visits during the planning study, video-recordings of IBL courses, a review of other college course observation protocols, and the literature on question-asking. Interpretation of the results is enhanced by comparing them with other types of data, such as student and faculty reports from surveys and interviews. In total, we analyzed nearly 300 hours of classroom observation data gathered from 42 separate sections, including 31 IBL and 11 non-IBL sections of 18 different courses on three campuses.

Our findings from classroom observation data are discussed in Chapter 2 and their linkages to student learning are reported in Chapter 4. Details of the observation sample, protocol and analysis methods are described in Appendix A2.

1.5.2 Student Surveys

Surveys are a flexible and powerful research tool that can be applied across all types of courses to yield quantitative data with good statistical power. We administered pre-surveys at the start and post-surveys at the end of IBL courses or, sometimes, multi-term sequences. With a large survey data set, we can compare patterns across different types of courses and contrast student groups by gender, race, ethnicity, college year, or academic background.

Our survey included three broad sections, or "instruments," that separately examined students' learning gains and course experiences, their attitudes and beliefs about mathematics, and their personal information. Different survey items required multiple-choice, numerical, or open-ended text responses. The learning gains instrument, SALG-M, was adapted from the existing SALG (Student Assessment of their Learning Gains) instrument, developed by our research group (Seymour, Wiese, Hunter & Daffinrud, 2000; SALG web site, n.d.) and used in a previous large-scale assessment (Weston, Thiry & Seymour, 2007). Items, or survey questions, on this instrument take the general form, "To what extent did you gain (learning outcome X)?" and "To what extent did (activity Y) help you learn?" This instrument was a post-test only, because students cannot report their learning gains, or growth, until they have completed the course.

The attitudinal instrument addressed aspects of students' attitudes and perceptions: confidence about doing mathematics; interest in mathematics as a field of study and as a personal interest; goals and motivations; problem-solving strategies; beliefs about mathematics learning; and beliefs about proof. Items on beliefs about teaching and learning mathematics were of particular interest for teacher preparation students.

Finally, the personal information instrument covered basic demographic variables, information about students' academic background (e.g. class status, self-reported GPA) and their prior experience with mathematics. Such data can be used to analyze survey answers by student group. It's also important to compare the initial populations of students enrolling in IBL and non-IBL courses, to ensure that any observed outcomes are not due to incoming differences. In undergraduate courses, the possibility of self-selection is high; in some courses, students were also institutionally selected by preferential invitation or advising into or out of IBL sections. Students constructed an identifier that enabled us to match pre- and post-survey responses without revealing individual identities. In some cases, we could match this identifier with other data, such as test scores.

Overall, we obtained a total of 1105 SALG-M responses, and 754 matched pre- and post-survey responses from 57 IBL and 18 non-IBL course sections. Various subsets of these are used for the analyses presented in Chapters 3 and 4. Details of the survey design, samples, and analysis methods are given in Appendix A3.

1.5.3 Tests of Mathematical Thinking and Learning

Tests are used to probe what students have learned or how they understand specific mathematical concepts. To be useful in research, direct assessments of student learning must be broad enough to apply across multiple courses, aligned to learning goals shared by IBL and non-IBL instructors, and high in content validity—that is, scores are seen by mathematicians as measuring something that is relevant and meaningful. Pre/post-test formats measure actual change over time, not just prior knowledge. For interpreting test scores, a comparative sample from the same course taught in non-IBL form is ideal. Ideally, tests are given during class, or to randomly selected students, so that study samples are representative of all students, not just volunteers.

We considered a variety of models for directly assessing student mathematical learning and explored with campus participants the opportunities for each:

a) <u>Standardized exams</u> developed by outside evaluators or researchers

b) <u>Common exam items</u> on midterm or final course tests, developed and given by instructors of individual IBL and non-IBL sections on the same campus

c) <u>Oral exams</u> by an expert panel of mathematicians, using a common set of questions and scoring rubric

d) A "<u>proof portfolio</u>" of authentic student work on important or challenging proofs. Expert panels would evaluate student solutions, then we would select some solutions for different students to evaluate and score students' evaluation of the proof in comparison with the expert evaluations.

Method (a) was of high interest, but we located no standardized tests that applied to the advanced classes (number theory, analysis, discrete math) that offered the most study opportunities. The recently developed Calculus Concept Inventory (Epstein, 2007; Epstein, Everson & Rhea, 2009) fit too few IBL courses at the study sites to be of use. For students taking mathematics courses for teacher preparation, we used parts of the *Learning Mathematics for Teaching* test (Hill, Schilling & Ball, 2004; Hill, Rowan & Ball, 2005). This test was developed to assess learning of practicing ("in-service") teachers from professional development workshops, rather than teacher candidates ("pre-service teachers") from undergraduate courses. Nonetheless, we found that the items were suitable for pre-service teachers and effectively discriminated levels of conceptual mastery. No comparison samples were available for these courses. In all, we obtained 109 sets of matched pre/post tests from three groups of pre-service teachers on two campuses, each of which took a multi-term course sequence targeted to them. The results from the LMT test are given in Chapter 5, with methodological details in Appendix A5.

Method (b), common exam items, was appealing in its simplicity. It is easily integrated into typical courses on a small scale. We attempted to engage pairs of IBL and non-IBL instructors in small assessment experiments to develop and administer exam questions that both believed would measure meaningful learning from their courses. We took care to promise anonymity and to avoid any suggestion of a "horse race" that would yield a winner and a loser. But we met with numerous barriers, including campus leaders' unwillingness to contact instructors about this option, instructors' concerns about "teaching to the test," a culture of last-minute test-writing, and skepticism about whether the results would be meaningful. Instructors raised concerns about the time required to develop common exam questions, yet were unwilling to consider exam questions that we proposed to develop with a group of TAs experienced in the same course. Ultimately we did not conduct any common-item experiments.

Method (c), oral exams, was based on a study by Wright et al. (1998). Given modest interest among study participants and the high logistical complexity of this design, we did not pursue this method.

A version of Method (d), the proof portfolio, was pursued. Rather than collecting and evaluating original student work, we presented students with a set of nine mathematical arguments taken from a research study by Weber (2009). We asked them to evaluate the arguments on several points: their understanding of the argument; its persuasiveness to them; and the level of explanatory power that it held, in their view. We also asked students to categorize the argument as a rigorous proof, a non-rigorous proof, or "not a proof." We piloted this approach in problem-solving interviews, where a student examined each problem and then explained his or her reasoning aloud to the interviewer. To shorten the time requirement and thus increase the sample size, we converted the test to a written form and gave the test to another sample. No instructors in sections with comparison groups opted to give the test in class, so both the interview and paper test samples were solicited from volunteers who had previously taken IBL or non-IBL sections of a number theory course at one campus. One IBL instructor at a second campus gave the test to his students as a pre/post test; no comparison section was available for that course.

In all, we obtained "proof test" data from 42 IBL and 35 non-IBL students (post only) at one campus, and pre/post tests (20 pre, 14 post) at a second campus. Results from the proof test are presented in Chapter 5, and further methodological detail is given in Appendix A5.

1.5.4 Academic Records

Grades are a traditional means of measuring academic achievement. How grades are assigned varies among instructors, but they are still widely assumed to hold meaning across academic settings. During our planning study, many IBL instructors offered hypotheses about student grades. Most commonly, they thought students' grades in the IBL course itself would not be comparable to those in a non-IBL course, because the nature of the graded work differed so much. However, they predicted, IBL students would carry improved learning habits and analytical skills to subsequent courses and outperform their non-IBL peers there. Instructors also suggested that IBL students would maintain interest in mathematics, perhaps taking more math courses or completing a math major in higher numbers. To evaluate these hypotheses, we examined academic records for patterns in grades, course selection, and academic majors.

We obtained academic records from three campuses[5] for all students enrolled in IBL and non-IBL sections of certain courses offered in specified academic terms. Our targets were core IBL courses that had a well-established history, sizable enrollments, and available comparison groups. In two cases, the comparison group was composed of students from non-IBL courses taught in the same general time period. At one campus with no simultaneous comparison sections, we experimented instead with a historical comparison group from the same course taught prior to the establishment of the IBL Center there. In each case, we chose course sections taught long enough ago so that then-enrolled students had likely completed their degree by the time of our data request, providing a fair comparison of "courses taken" and "grades obtained" up to graduation.

Using enrollment in the targeted courses to define the study population, each campus' registrar or institutional research office provided enrollment and grade data on all mathematics courses these students had taken before, during, and after the IBL course of interest ("target" course), and data on majors declared or dropped. Because institutional records varied widely in format, raw data were converted into standardized variables that counted course enrollments in particular sets of later math courses and computed GPAs for these courses.

To compare patterns across student groups and to test for selective enrollment in IBL or non-IBL sections, we also requested demographic data and admissions data such as high school grade point average (HS GPA), calculus Advanced Placement (AP) score, and scores on standardized tests of college readiness (SAT or ACT). These data were also converted to standard formats. In one case, selection into the small IBL section versus the large, non-IBL, lecture section was very strong, so we used demographic and admissions data to select a small comparison sample that closely matched the IBL student population. In other cases, direct comparison of IBL and non-IBL sections was possible. We did not conduct a transcript analysis for the targeted pre-service teacher courses because these students are unlikely to take additional mathematics courses beyond their program requirements, and because no non-IBL comparison group existed.

---

[5] One campus did not respond to our requests for academic records data.

Ultimately, we analyzed data from three courses on two campuses, totaling 3212 student records. Results of our statistical analyses are presented in Chapter 6, and full details of the samples, constructed variables, and analysis methods are given in Appendix A6.

1.5.5 Interviews

Interviews provide detailed insight into the experiences and views of students and instructors. In this study, interview data help to confirm observations from other sub-studies and to clarify classroom processes that account for the learning outcomes measured by other means. Because interviews are open-ended and allow respondents to spontaneously offer their own ideas, they uncover new issues that one might not think to include in a survey. While interview data often resonate with readers because they resemble personal and perhaps familiar stories, they differ from anecdotes. Interview samples are carefully structured, and interview questions are composed to systematically elicit commentary on specific research issues. Researchers dissect the transcript using a rigorous analytical process to examine and categorize every statement or passage. Sometimes we count the frequency of comments on specific issues to see if certain experiences or opinions are common or rare among different groups. We may also compare the perspectives of students and instructors on the same classroom phenomenon.

We selected specific IBL "core" courses for each IBL Center and created structured samples that took into account students' gender, major, and class year. Because students of certain racial and ethnic backgrounds are underrepresented in mathematics, both nationally and at these campuses, we invited students of color to interview whenever we could identify them. We invited students to schedule an interview during our two- to three-day visit on their campus. The interview questions addressed students' experiences in their IBL course, what they did or did not learn in the course, their academic background and future intentions, and their advice to other students and instructors taking or teaching an IBL course. Students from the same course and demographic group were sometimes combined into focus groups of two or three. In all, we interviewed 68 students representing all four campuses in 41 individual or focus group sessions.

We also interviewed instructors, including both faculty and graduate students. We contacted all lead instructors and teaching assistants who had taught an IBL course in the past three years; over three-quarters granted an interview in person or by telephone. Interview questions addressed instructors' experiences teaching and learning to teach IBL courses, their observations of student gains, and the relationship of IBL teaching to their own career development. We asked their views of their departmental and institutional context, and their advice to other instructors teaching an IBL course for the first time. We obtained interview data from 23 faculty instructors and 20 graduate student instructors representing all four campuses.

Both student and instructor interviews were transcribed verbatim and analyzed by qualitative coding. Each text passage was examined and assigned one or more labels marking the ideas and themes it raised. Passages sharing common themes were gathered under the same label, or code, so that they could be analyzed as a group, sorted by sub-theme, and, if useful, counted.

Chapter 7 discusses findings from student interview data on student learning outcomes and learning processes. Chapter 8 describes findings from instructor interviews. Details of the interview samples, protocols, and analysis methods are provided in Appendix 7A.

1.6 Terminology Used in the Report

To assist our primary audience, and for reference, we include brief definitions of a few key terms that are used repeatedly throughout the report.

1.6.1 Terms Used in Describing Study Populations

IBL course: a course designated by its campus leader as using inquiry-based learning methods, and typically supported by Center funds. Teaching and learning approaches in these courses varied, and some non-IBL courses used student-centered methods as well. However, observation data confirm that most IBL-designated courses used student-centered methods substantially more than most of their non-IBL counterparts (Chapter 2).

Math-track students: our term for students taking challenging mathematics courses, from first-year to advanced level. In this study, most were math, science, or engineering majors; some were also pursuing a teaching certificate for high school mathematics.

Pre-service teachers: undergraduates preparing to teach K-12 school but not yet certified. Elementary (K-8) certification is typically broad and requires only a few mathematics courses. In this project, some IBL courses were targeted to this population. Secondary (9-12) certification students typically graduate with both a math major and specialized preparation for teaching. In this study, one IBL course sequence was targeted to pre-service secondary teachers, and some were also enrolled in advanced IBL courses for math majors.

Instructors: all those involved in teaching a course, including faculty and graduate students. In this study, faculty were permanent tenure-track faculty (pre- or post-tenure), permanent non-tenure-track faculty (called instructors or lecturers at various institutions), or temporary faculty in postdoctoral positions. Graduate students played teaching roles that included lead instructor, teaching assistant (TA), or college fellow, a novice teaching assistant who assisted with grading, class time, and office hours as preparation for future TA roles.

Target course: IBL or comparative course that is the intervention of focus and thus defines a study population for the academic records study (Chapter 6). We examined the courses that students chose following their participation in an IBL (or comparative) target course, and their grades in those later courses. We also checked on students' course choices before the target course, to ensure that any differences observed were not due to previous background.

1.6.2 Terms Used in Describing Study Results

Gains: growth or positive change of some kind that may be attributed to an experience or treatment (e.g. a math course). In this study, we include gains in confidence, learning particular concepts or relationships among concepts, understanding the nature of the discipline, skills or abilities to do things, appreciation, interest, and so on.

In this report, we discuss gains both quantitatively and qualitatively. On the learning gains portion of the SALG-M survey, students reported their own gains on a scale of 1 (little or no gain) to 5 (great gain). In analyzing interviews and open-ended survey comments, we categorize das gains all student comments about what they learned or took away from a course. We similarly coded instructors' observations of their students' gains (or lack of them). Occasionally we speak of negative gains or losses, meaning a negative change due to the experience. We use the phrase test score gains to refer to changes in test score from start to end of a course, e.g. differences between pre- and post-scores on a test (Section 5.2.2).

Statistical significance: a standard assessment of the trustworthiness of a measurement for some property that naturally or randomly varies. A "statistically significant" result is less likely to be solely due to chance. For example, a significant difference between two groups is more likely to represent a real difference in two groups. If there were *no* true difference, the probability of seeing *by chance alone* a difference at least as big as the difference you *actually measured* is less than 5% (p value < 0.05)

Statistical significance does not ensure that the result is meaningful, but it does generally mean that researchers will attempt to interpret the result. Statistical significance depends on both the size of the difference and the size of the sample, and is often reported in terms of the p-value. For readers unfamiliar with these terms, we recommend the nontechnical discussion, using medical examples, at http://cancerguide.org/significance.html

p-value: the probability that a specific result is due to chance or random variation alone. For example, a result reported with $p<0.05$ means that there is less than 5% chance that the result is due to chance alone. A result with $p<0.01$ means that there is less than 1% chance of this. So a smaller p-value means that it is less probable that the result is due to chance alone. In this report, we report selected p-values in tables and text. Our discussion emphasizes statistically significant results, but in some cases we do discuss repeated patterns of results that have p-values above 0.05. See also *statistical significance*.

Effect size (symbol *d*): a standardized way to report the magnitude of the effect of some treatment (e.g. the effect of taking a particular math class). Effect size is expressed in terms of standard deviations, and is therefore independent of the units of measurement, enabling different kinds of measures to be compared. The larger the effect size, the bigger the difference the treatment appears to make. For example, in Section 3.4.4 we compare the effect sizes of IBL (or non-IBL) courses on students' motivation and approaches to learning. Interpreting effect sizes depends on the measures used and the samples studied, but educational researchers often interpret values of 0.2-0.3 as "small," 0.3-0.7 as "medium," and 0.7 or above as "large" effects. See also *statistical significance*.

Variance: a statistical measure of the variation in a study sample. We use variance to test how much certain factors may explain or account for a particular outcome. For example, in Section 3.3.1 we examine whether variations in students' course *experiences* account for the differences in their learning *gains*. The results are expressed as the percentage of the variance in gains that is explained by each of several different course experiences. Since there is always some random or unexplained variation (e.g., due to factors not studied), explaining 10% or more of a variance is considered a strong contribution of a single factor. Preferably, a model with multiple factors will explain 40% of the variance or more.

Correlation (symbol *r*): a statistical measure of the relationship between two variables, varying from -1 to +1. A large absolute value of the correlation (positive or negative) means the variables are strongly related; a value close to zero means they are related only weakly. For example, in Chapter 4, the correlations between classroom observation variables and student learning outcomes indicates the extent to which these independent variables are related, and thus suggests (though does not prove) that the relation may be causal.

1.7 Acknowledgments

We have many people to thank for their help with this project. We are especially grateful to all the students and instructors who participated in one or more parts of the study. Without their effort, interest, and candor, we could learn nothing. We also thank the following individuals for specific assistance in providing information, ideas, or resources, making arrangements, reviewing materials, or gathering data:

- IBL Center leaders: John Boller, Bill Jacob, Ted Odell, Paul Sally, Ralf Spatzier, Mike Starbird
- IBL Center colleagues: Elisa Bass, Chris Betz, Bill Breslin, Mort Brown, Mark Daniels, Doreen Fussman, Monica Guzman, Vilma Mesa, Katie Pelletier, Cyntreva Paige, Jessica Taylor, Stephanie Walthes
- Student records data: Andy Cameron, Michele Keeler, Steven Velasco
- Classroom observation data: Monica Guzman, Sarah Hough, Joseph Hunt, Dave Jensen, Amy Hiumin Kao, Jess Kennedy, Kyunghee Moon, Mark Rothlisberger, Puttoei Talawat, Tim Whittemore
- Proof interviews: Steven Greenstein, Dave Jensen
- Literature reviews: Sarah Hough
- Boulder colleagues: Clint Coburn, Rebecca Crane, Heather Thiry, Tim Weston
- Boulder student workers: Justin Adkins, Clint Coburn, Sam Garrett, Sarah Kleinman, Nathan Mason, Brandon Mells, Michelle Six
- Educational Advancement Foundation: Deborah Cole, Ron Douglas, Norma Flores, Albert Lewis, Harry Lucas, Marty Reiner
- Elsewhere: Bill Calhoun, Jerome Epstein, Tom Judson, Brian Katz, Kyle Peterson, Geoffrey Phelps, Keith Weber, Stan Yoshinobu

This work was supported by grants from the Educational Advancement Foundation based in Austin, Texas. All conclusions are those of the authors and do not necessarily reflect the views of the Foundation. We thank Susan Millar for initiating this project and for generously sharing her preliminary design work with us.

1.8 References Cited

Adamson, S. L., Banks, D., Burtch, M., Cox, F., Judson, E., & Turley, J. B. (2003). Reformed undergraduate instruction and its subsequent impact on secondary school teaching practice and student achievement. *Journal of Research in Science Teaching*, *40*(10), 939–957.

Alrø, H., & Skovsmose, O. (2002). *Dialogue and learning in mathematics education: Intention, reflection, critique*. Boston: Kluwer Academic Publishers.

Alsardary, S., & Blumberg, P. (2009). Interactive, learner-centered methods of teaching mathematics. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies, 19*(4), 401-416.

Ball, D. L., Hill, H. C., & Bass, H. (2005, Fall). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator*, *29*, 14-22.

Banton, M. L. (2002). Using an undergraduate geometry course to challenge pre-service teachers' notions of discourse. *Journal of Mathematics Teacher Education, 5*, 117-152.

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, *107*(5), 1860-1863.

Ben-Chaim, D., Keret, Y., & Ilany, B. S. (2007). Designing and implementing authentic investigative proportional reasoning tasks: The impact on pre-service teachers' content and pedagogical knowledge and attitudes. *Journal of Mathematics Teacher Education, 10,* 333-340.

Bolden, D. S., Harries, T. V., & Newton, D. P. (2010). Preservice primary teachers' conceptions of creativity in mathematics. *Educational Studies in Mathematics*, *73*, 143-157.

Bonsangue, M. (1991). Achievement effects of collaborative learning in introductory statistics: A time series residual analysis. Paper presented at the Joint Annual Meeting of the Mathematical Association of America and The American Mathematical Society, San Francisco, CA.

Bonsangue, M. (1994). An efficacy study of the calculus workshop model. In Dubinsky, E., Schoenfeld, A. H., & Kaput, J. J. (Eds.), *Research in Collegiate Mathematics Education I*, (pp 117-137). Providence, RI: American Mathematical Society in cooperation with Mathematical Association of America .

Bookman, J., & Friedman, C. P. (1994). A comparison of the problem solving performance of students in lab based and traditional calculus. In Dubinsky, E., Schoenfeld, A. H., &Kaput, J. J. (Eds.), *Research in Collegiate Mathematics Education I*, (pp 101-116). Providence, RI: American Mathematical Society in cooperation with Mathematical Association of America.

Bookman, J., & Friedman, C. P.(1999). The evaluation of Project CALC at Duke University, 1989-1994. Mathematical Association of America, MAA Notes #49. Retrieved 2/2/2011 from http://www.maa.org/saum/maanotes49/toc.html

Bruner, J. S. (1961). The act of discovery. *Harvard Education Review*, *31*(1), 21-32.

Buch, N. J., & Wolff, T. F. (2000). Classroom teaching through inquiry. *Journal of Professional Issues in Engineering Education and Practice,* (July), 105-109.

Buck, R. E. (2006). Conjecturing. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies, 6*(2), 97-104.

Chapman, O. (2007). Facilitating preservice teachers' development of mathematics knowledge for teaching arithmetic operations. *Journal of Mathematics Teacher Education, 10,* 341-349.

Christian Smith, J. (2006). A sense-making approach to proof: Strategies of students in traditional and problem-based number theory courses. *Journal of Mathematical Behavior, 25*(1), 73-90.

Cobb, P., Yackel, E., & McCain, K. (Eds.)(2000). *Symbolizing and communicating in mathematics classrooms: Perspectives on discourse, tools, and instructional design.* Mahwah, NJ: Lawrence Erlbaum Associates.

Colburn, W. (1826). *Arithmetic upon the inductive method of instruction*. Cited in Jacob, W. (no date). The history and promise of IBL in math education. UCSB Department of Mathematics, Center for Mathematical Inquiry. Retrieved 4/27/2011 from http://math.ucsb.edu/department/cmi/IBL_History.html

Committee on Equal Opportunities in Science and Engineering (CEOSE) (2005). *Broadening Participation in America's Science and Engineering Workforce; The 1994-2003 decennial and 2004 biennial reports to Congress.* B. Hartline, report chair. CEOSE 04-02. Retrieved 4/15/2011 from http://www.nsf.gov/od/oia/activities/ceose/

Committee on Science, Engineering and Public Policy (COSEPUP) (2007). *Rising Above The Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. Committee on Prospering in the Global Economy of the 21st Century. Washington DC: National Academy Press.

Conference Board of Mathematical Sciences (2001). *The Mathematical Education of Teachers.* Providence RI and Washington DC: American Mathematical Society and Mathematical

Association of America. Retrieved 11/10/2010 from http://www.cbmsweb.org/MET_Document/index.htm

Dewey, J. (1938). *Experience and education*. New York: Collier.

Duch, B. J., Groh, S. E., & Allen, D. E. (2001). *The power of problem-based learning.* Sterling, VA: Stylus.

Dumitraşcu, D. (2009) Integration of guided discovery in the teaching of real analysis. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies, 19*(4), 370-380.

Educational Advancement Foundation (EAF) (2007, February). Request for Proposal on Assessment & Evaluation Center for Transforming American Mathematics Education, Inquiry-Based Learning Project in Mathematics.

Enemaker, C. (1995). A problem-solving based mathematics course and elementary teachers' beliefs. *School Science and Mathematics*, 75-83.

Epstein, J. (2007). [RUME] Calculus Concept Inventory. E-mail message to Research on Undergraduate Mathematics Education mailing list, November 28, 2007. Retrieved 1/27/2011 from http://betterfilecabinet.com/pipermail/rume_betterfilecabinet.com/2007-November/001112.html

Epstein, J., Everson, H., & Rhea, K. (2009). The Calculus Concept Inventory—Measuring student understanding in calculus. Preprint received as personal communication from J. Epstein to S. Laursen, November 29, 2009.

Freudenthal, H. (1991). *Revisiting mathematics education*. Dordrecht, The Netherlands: Kluwer.

Ganter, S. L., & Jirovtek, M., R. (2000). The need for evaluation in the calculus reform movement: A comparison of two calculus teaching methods. In Dubinsky, E., Schoenfeld, A. H., & Kaput, J. J. (Eds.), *Research in Collegiate Mathematics Education IV* (pp 16-41). Providence, RI: American Mathematical Society in cooperation with Mathematical Association of America.

Gillies, R. M. (2007). *Cooperative learning: Integrating theory and practice*. Thousand Oaks, CA: Sage Publications.

Goodsell, A., Mahler, M., Tinto, V., Smith, B. L., & MacGregor, J. (Eds.) (1992). *Collaborative learning: A sourcebook for higher education.* University Park, PA: National Center on Postsecondary Teaching, Learning and Assessment.

Gravemeijer, K. P. E. (1994). *Developing realistic mathematics education*. Utrecht, The Netherlands: Cd-B Press.

Hart, C. H. (2002). Preservice teachers' beliefs and practices after participating in an integrated content/methods course. *School Science and Mathematics, 1*, 4-14.

Herzig, A., & Kung, D. T. (2003). Cooperative learning in calculus reform: What have we learned? In Selden, A., Dubinsky, E., Harel, G., & Hitt, F. (Eds.), *Research in Collegiate Mathematics Education V* (pp. 30-55). Providence, RI: American Mathematical Society in cooperation with Mathematical Association of America.

Hill, H. C., & Ball, D. (2004). Learning mathematics for teaching: Results from California's mathematics professional development institutes. *Journal for Research in Mathematics Education*, 35(5), 330-351.

Hill, H., C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, *39*(4), 372-400.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371-406.

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal*, *105*(1), 11-30.

Hough, S. (2010a). The effects of the use of inquiry-based learning in undergraduate mathematics on student outcomes: A review of the literature. [Report prepared for Ethnography & Evaluation Research, University of Colorado at Boulder.]

Hough, S. (2010b). The effects of the use of inquiry-based learning in undergraduate mathematics on prospective teacher outcomes: A review of the literature. [Report prepared for Ethnography & Evaluation Research, University of Colorado at Boulder.]

Hough, S., Jacob, B., Mendoza, M., & Moon, K. (2011). The development of prospective elementary teachers' understandings of children's mathematics. Paper accepted to the Annual Meeting of the American Educational Research Association. April 8-12, New Orleans, Louisiana.

Hurley, J. F., Koehn, U., & Ganter, S. L. (1999). Effects of calculus reform: Local and national. *American Mathematical Monthly*, *106*(9), 800–811.

Jacob, W. (no date). The history and promise of IBL in math education. UCSB Department of Mathematics, Center for Mathematical Inquiry. Retrieved 4/27/2011 from http://math.ucsb.edu/department/cmi/IBL_History.html

Jensen, J. (2006). Surprising effects of inquiry based learning. *Texas College Mathematics Journal, 3*(1), 1-9.

Johnson, D., Johnson, R., & Smith, K. (1998). *Active learning: Cooperation in the college classroom,* 2nd ed. Edina, MN: Interaction Book Co.

Ju, M., & Kwon, O. N. (2007). Ways of talking and ways of positioning: Students' beliefs in an inquiry-oriented differential class. *Journal of Mathematical Behavior, 26*, 267-280.

King, A. (2002). Structuring peer interaction to promote high-level cognitive processing. *Theory into Practice, 41*, 33-40.

Kwon, O. N., Rasmussen, C., & Allen, K. (2005). Students' retention of mathematical knowledge and skills in differential equations. *School Science and Mathematics, 105*(5), 1-13.

Larsen, S. (2009). Reinventing the concepts of group and isomorphism: The case of Jessica and Sandra. *Journal of Mathematical Behavior, 28*(2-3), 119-137.

Larsen, S., & Zandieh, M. (2008). Proof and refutations in the undergraduate mathematics classroom. *Educational Studies in Mathematics, 67*, 207-216.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

Lawrenz, F., Huffman, D., & Gravely, A. (2007). Impact of the Collaboratives for Excellence in Teacher Preparation. *Journal of Research in Science Teaching*, *44*(9), 1247-1415.

Leitzel, J. R. C. (ed). (1991). *A call for change: Recommendations for the mathematical preparation of teachers of mathematics.* Washington, DC: Mathematical Association of America.

Lerman, S. (2000). The social turn in mathematics education research. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 19-44). Palo Alto, CA: Greenwood.

Lo, J. J., Grant, T. J., & Flowers, J. (2007). Challenges in deepening prospective teachers' understanding of multiplication through justification. *Journal of Mathematics Teacher Education 11*, 5-11.

Lubinski, C. A., & Otto, D. A. (2010). Preparing K-8 preservice teachers in a content course for standards-based mathematics pedagogy. *School Science and Mathematics*, 336-350.

Manigault, S. (1997). *The book for math empowerment: Rethinking the subject of mathematics*. Stafford, VA: Godosan Publications.

Mathews, M. E., & Seaman, W. (2007, July). The effects of different undergraduate mathematics courses on the content knowledge and attitude towards mathematics of preservice elementary teachers. *Issues in the Undergraduate Preparation of School Teachers: The Journal* [Online] Vol. 1: Content Knowledge, (July). Retrieved 2/2/2011 from http://www.k-12prep.math.ttu.edu/journal/contentknowledge/volume.shtml

McCann, T. M, Johannessen, L. R., Kahn, E., & Smagorinsky, P. (Eds.) (2005). *Reflective teaching, reflective learning: How to develop critically engaged readers, writers, and speakers*. Portsmouth, NH: Heinemann.

McClain, K. (2003). Supporting preservice teachers' understanding of place value and multidigit arithmetic. *Mathematical Thinking and Learning, 5*(4), 281-306.

McDonald, M. A., Mathews, D. M., & Strobel, K. (2000). Understanding sequences: A tale of two objects. In Dubinsky, E., Schoenfeld, A. H., & Kaput, J. J. (Eds.), *Research in Collegiate Mathematics Education IV* (pp 77-102). Providence, RI: American Mathematical Society in cooperation with Mathematical Association of America.

McGinnis, J. R., Kramer, S., Shama, G., Graeber, A. P., Parker, C. A., & Watanabe, T. (2002). Undergraduates' attitudes and beliefs about subject matter and pedagogy measured periodically in a reform-based mathematics and science teacher preparation program. *Journal of Research in Science Teaching, 39*(8) 713-737.

McLoughlin, M. P. M. M. (2009). Inquiry-based learning: An educational reform based upon content-centered teaching. Paper presented at the Annual Meeting of the Mathematics Association of America, Washington, DC, January 5-8, 2009.

McLoughlin, M. P. M. M. (2008). Crossing the bridge to higher mathematics: Using a Modified Moore approach to assist students transitioning to higher mathematics. Paper presented at the Annual Meeting of the Mathematics Association of America, San Diego, CA, January 6, 2006.

Meel, D. E. (1998). Honors students' calculus understandings: Comparing *Calculus & Mathematica* and traditional calculus students. In Dubinsky, E., Schoenfeld, A. H., & Kaput, J. J. (Eds.), *Research in Collegiate Mathematics Education III* (pp.163-215). Providence RI: American Mathematical Society in cooperation with Mathematical Association of America.

Members of the 2005 "Rising Above the Gathering Storm" Committee (2010). *Rising Above the Gathering Storm, revisited: Rapidly approaching Category 5*. Prepared for the Presidents of the National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. Washington, DC: National Academies Press.

Mezirow, J. (1991). *Transformative dimensions of adult learning.* San Francisco, CA: Jossey-Bass.

Moon, J. A. (2004). *A handbook of reflective and experiential learning: Theory and practice.* New York: RoutledgeFalmer.

Narli, S., & Baser, N.(2008). Cantorian set theory and teaching prospective teachers. *International Journal of Environmental & Science Education*, 3(2), 99-107.

National Council of Teachers of Mathematics (NCTM) (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: NCTM.

National Council of Teachers of Mathematics (NCTM) (1991). *Professional standards for teaching mathematics.* Reston, VA: NCTM.

National Research Council (NRC) (1989). *Everybody counts: A report to the nation on the future of mathematics education.* Washington, DC: National Academy Press.

National Science Board (2010). *Science and Engineering Indicators 2010*. Arlington, VA: National Science Foundation (NSB 10-01). Retrieved 1/27/2011 from http://www.nsf.gov/statistics/seind10/start.htm.

Owens, K., Perry, B., Conroy, J., Geoghagen, N. & Howe, P. (1998). Responsiveness and effective processes in the interactive construction of understanding in mathematics. *Educational Studies in Mathematics, 35*, 105-127.

Paulos, J. A. (1988). *Innumeracy: Mathematical illiteracy and its consequences*. New York: Hill & Wang.

Paulos, J. A. (1995). *A mathematician reads the newspaper.* New York: Basic Books.

Penn, H. (1994). Comparisons of test scores in Calculus I at the Naval Academy. *Focus on Calculus, a Newsletter for the Calculus Consortium Based at Harvard University*. Wiley (6), 6-7.

Perkins, D. N., & Tishman, S. (2001). Dispositional aspects of intelligence. In Collis, J. M., & Messick, S. (Eds.), *Intelligence and personality: Bridging the gap in theory and measurement* (pp. 233-257). Mahwah, NJ: Lawrence Erlbaum.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*(3), 223-231.

Prince, M, & Felder, R. (2007). The many facets of inductive teaching and learning. *Journal of College Science Teaching, 36*(5), 14-20.

RAND Mathematics Study Panel (2003). *Mathematical proficiency for all students: Towards a strategic research and development program in mathematics education*. Santa Monica, CA: RAND.

Rasmussen, C., & Blumenfield, H. (2007). Reinventing solutions to systems of linear differential equations: A case of emergent models involving analytic expressions. *Journal of Mathematical Behavior, 26*(3), 195-210.

Rasmussen, C., Kwon, O. N., Allen, K., Marrongelle, K., & Burtch, M. (2006). Capitalizing on advances in mathematics and K-12 mathematics education in undergraduate mathematics: An inquiry-oriented approach to differential equations. *Asia Pacific Education Review, 7*(1), 85-93.

Roddick, C. (2003). Calculus reform and traditional students' use of calculus in an engineering mechanics course. In Selden, A., Dubinsky, E., Harel, G., & Hitt, F. (Eds.), *Research in Collegiate Mathematics Education V*, (pp 56-78). Providence, RI: American Mathematical Society.

SALG home page (n.d.). Student Assessment of their Learning Gains. About SALG. Retrieved 1/31/2011 from http://www.salgsite.org/about

Savin-Baden, M., & Major, C. H. (2004). Foundations of problem-based learning. Maidenhead, UK: Open University Press.

Schwingendorf, K. E., McCabe, G. P., & Kuhn, J. (2000). A longitudinal study of the C^4L calculus reform program: Comparisons of C^4L and traditional students. In Dubinsky, E., Schoenfeld, A., & Kaut, K. (Eds.), *Research in Collegiate Mathematics Education IV* (pp 63-102). Providence, RI: American Mathematical Society.

Seymour, E., Wiese, D., Hunter, A.-B., & Daffinrud, S. M. (2000, March). Creating a better mousetrap: On-line student assessment of their learning gains. Presented at the 219th National Meeting of the American Chemical Society, San Francisco, CA. Retrieved 1/25/2011 from http://salgsite.org/docs/SALGPaperPresentationAtACS.pdf

Shriki, A. (2010). Working like real mathematicians: Developing prospective elementary teachers awareness of mathematical creativity through generating new concepts. *Educational Studies in Mathematics*, *73*, 159-179.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(5), 4-14.

Smith III, J., & Star, J. (2007). Expanding the notion of impact of K-12 standards-based mathematics and reform calculus programs. *Journal for Research in Mathematics Education*, *38(1)*, 3-34.

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research*, 69(1), 21-51.

Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education*, *40*(3), 314-352.

Szydlik, J. E., Szydlik, S. D., & Benson, S. R. (2003). Exploring changes in pre-service elementary teachers' mathematical beliefs. *Journal of Mathematics Teacher Education, 6,* 253-27.

Treisman, U. (1986). A study of the mathematics performance of black students at the University of California, Berkeley. (Doctoral dissertation, University of California, Berkeley, 1985). *Dissertation Abstracts International*, 47, 1641-A.

Tucker, A. (1995). *Assessing calculus reform efforts: A report to the community*. The Mathematical Association of America.

Urion, D. K., & Neil, D. (1992). Student achievement in small-group instruction versus teacher-centered instruction in mathematics. *PRIMUS: Problems, Resources and Issues in Mathematics Undergraduate Studies, 2*(3), 257-64.

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

Wagner, J. F., Speer. N. M., & Rossa, B. (2007). Beyond mathematical content knowledge: mathematician's knowledge needed for teaching an inquiry-oriented differential equations course. *The Journal of Mathematical Behavior, 26*, 247-266.

Weber, K. (2009). Proving is not convincing. Conference on Research in Undergraduate Mathematics Education, Raleigh, NC, February 26-March 1, 2009.

Weston, T. J., Thiry, H., & Seymour, E. (2007). Comparing three approaches to science teaching reform and their implementation: Results from the *Science Education for New Civic Engagement and Responsibilities* project. Unpublished report.

Wilkins, J. L. M., & Brand, B. R, (2004). Change in preservice teachers' beliefs: An evaluation of a methods course. *School Science and Mathematics*, 104(5), 226-232.

Williams, C. G. (1998). Using concept maps to assess conceptual knowledge of function. *Journal for Research in Mathematics Education, 29*(4), 414-421.

Wright, J. C., Millar, S. B., Kosciuk, S. A., Penberthy, D. L., Williams, P. H., & Wampold, B. E. (1998). A novel strategy for assessing the effects of curriculum reform on student competence. *Journal of Chemical Education, 85*(8), 986-992.

Yackel, E., & Rasmussen, C. (2002). Beliefs and norms in the mathematics classroom. In Leder, G., Pehkonen, E., & Torner, G. (Eds.), *Beliefs: A hidden variable in mathematics education?* (313-344). Dordrecht: Kluwer Academic Publishers.

Yackel, E., Rasmussen, C. , & King, K. (2000). Sociomathematical norms in an advanced undergraduate mathematics course. *Journal of Mathematical Behavior, 19*, 275-287.

Yoo, S., & Christian Smith, J. (2007). Differences between mathematics majors' view of mathematical proof after lecture-based and problem-based instruction. In Lamberg, T., & Wiest, L. R. (Eds.). (2007). Proceedings of the 29th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Stateline (Lake Tahoe), NV: University of Nevada, Reno.

Zacharie, M. (2009). Why college or university students hate proof in mathematics? *Journal of Mathematics and Statistics, 5*(1), 32-41.

Zimmerman, B. J., & Schunk, D. H. (2001). *Self-regulated learning and academic achievement: Theoretical perspectives.* Mahwah, NJ: Lawrence Erlbaum Associates.

Chapter 2: Findings from Classroom Observation

2.1 Overview of the Observation Study

Data from classroom observation enable us to document classroom practices and interactions in IBL and non-IBL teaching practices, and to link these practices to student outcomes. We sought to address the following questions:

- How are classrooms designated "IBL" alike or different from comparative classrooms?
- What practices and features are commonly applied in IBL courses, and what are the variations among them?
- What classroom features are seen in course sections where there are good student outcomes? How do these compare with classroom features of sections where outcomes are less positive?

The study methods are detailed in Appendix A2.[1] Briefly, we logged and categorized the instructional activity taking place during every minute of class and recorded who was in the leadership role—instructors or students, individuals or groups. We categorized all questions asked by students and instructors, the type of each question (cognitive level, other classroom function), and the instructional activity taking place when the question was asked. Observers also rated aspects of the classroom atmosphere and student/instructor interactions on a simple post-class survey, with written comments to support their ratings. Our observations focused on the inclusion and features of student-centered instructional activities, rather than on instructors' skill in implementing them. We analyzed data for 31 IBL course sections and 11 non-IBL sections of a subset of the same courses.

2.2 Differences in Classroom Practice between IBL and non-IBL Sections

Below we discuss some key differences between IBL and non-IBL sections, comparing mean values of observation variables that describe the nature of instructional activities, leadership roles, and question-asking behaviors. The statistical significance of differences in group means is given in the tables. Key findings are called out in the italicized headings of each sub-section.

2.2.1 More class time was spent on student-centered instructional activities in IBL classrooms.

As Table 2.1 (Section A) shows, student-centered instructional activities accounted for significant proportions of class time in IBL classrooms, but very small amounts of time in non-IBL classrooms. These might include student presentation (typically, a student at the board showing a proof or solution), small group work, and, in a few courses, computer work such as simulation activities. Whole-class discussion was also more significant in IBL classrooms. The time allocated to each of these activities varied widely; but on average, over 60% of total IBL class time was spent in these student-centered activities vs. less than 10% of non-IBL class time.

---

[1] http://www.colorado.edu/eer/research/documents/A2classroomObservationMethods.pdf

Table 2.1: Observed Instructional Activities in IBL (N=31) & Non-IBL (N=11) Classrooms

Observation variable	**mean for IBL sections**	**mean for non-IBL sections**	**Sig.**
A. Percentage of class time spent on instructional activities			
Student-centered	**62.4%**	**8.1%**	***
Student presentation	24.0%	2.1%	***
Discussion	12.1%	3.1%	***
Group work	24.2%	3.0%	***
Computer work	2.2%	0.0%	*
Instructor-centered	**27.0%**	**87.1%**	***
Lecture	19.2%	79.1%	***
Explanation	7.8%	8.0%	
Remainder	**10.6%**	**4.8%**	**
Class business	6.9%	4.0%	*
Other	3.8%	0.8%	**
B. Activity density (number of distinct episodes of activity per hour of class time)			
Student-centered	**4.54**	**0.60**	***
Student presentation	2.35	0.22	***
Discussion	1.11	0.30	**
Group work	0.99	0.08	***
Computer work	0.10	0.00	**
Instructor-centered	**2.65**	**1.91**	
Lecture	1.61	1.38	
Explanation	1.05	0.53	*
Remainder			
Class business	1.12	0.68	*
Other	0.30	0.08	*
Overall variety: Total number of episodes of distinct activity per hour	**8.61**	**3.28**	***
C. Length of episodes of instructional activities, minutes			
Student-centered			
Student presentation	7.2	5.8	
Discussion	8.5	6.9	
Group work	14.5	19.3	
Computer work	12.1	-	
Instructor-centered			
Lecture	9.2	43.2	***
Explanation	6.2	10.2	**
Remainder			
Class business	4.4	3.4	
Other	8.7	1.0	

*** p<0.001; ** p<0.01; * p<0.05

In contrast, in non-IBL sections, nearly 90% of class time was spent on instructor-centered activities, including both prepared lectures and spontaneous explanations by instructors. These accounted for about one quarter of IBL class time.

Class business seemed to require somewhat more time in an IBL class. This included time to (e.g.) get a group activity started, hand back papers, or (often) wait while a student wrote a proof or solution on the board.

2.2.2 Instructional activities were more varied in IBL classrooms.

One indicator of variety is the number of separate episodes of each type of instructional activity (Table 2.1, Section B). For example, a lecture that lasted the entire class constituted one episode. As the activity changed, new episodes of lecture, discussion, group work, etc., were recorded.

Overall, IBL classrooms showed a greater variety of activity, with more than twice as many distinct episodes of instructional activity per hour of class observed. This variety was largely due to the inclusion of student-centered activities (including presentation, group work, discussion, and computer work), which were over seven times more common in IBL classes.

Episodes of faculty activity were equally common in IBL and non-IBL classrooms—but were, on average, much shorter in duration in IBL classes (Table 2.1B/C). For example, lectures took place similarly often in both types of classes, but in IBL classes, lectures averaged 9 minutes in length, vs. 42 minutes in non-IBL classes. Short, spontaneous "explanations" were also more common in IBL classrooms, often (86%) by instructors but also by students. In addition to the deeper engagement possible from student-centered activities, the frequent change of activity in IBL classes may have helped to hold or revive students' attention.

2.2.3 Students took a greater leadership role in IBL classrooms.

We categorized each episode of instructional activity (presentation, group work, etc.) according to the leadership role of students (as individuals, groups, or the class as a whole) or by instructors. While this often corresponded to the nature of instructional activity itself (e.g. faculty commonly led a lecture), these two categorizations did not always match one-to-one. Table 2.2 shows the means for the observed leadership-related variables.

On average, students led IBL classroom activities for over ten times the total time they led in non-IBL classes. Group and whole-class leadership were particularly noteworthy, as these roles were nearly entirely absent in non-IBL classes. TAs also took a greater role in leading IBL classes, unlike the non-IBL classes observed. On average, instructors led non-IBL classes for 95% of time, vs. about 40% for IBL classes.

Likewise, episodes of student leadership were ten times more frequent in IBL classes. And, while instructors were in the lead role for *less time* during an IBL class, they took the lead *more often*. Part of this difference is due to TAs' greater participation: an IBL TA took an active leadership role once every two class hours (vs. once in 20 hours for non-IBL classes). In general, instructors contributed in short bursts rather than holding the floor at length.

Table 2.2: Observed Leadership Roles in IBL (N=31) & Non-IBL (N=11) Classrooms

Observation variable	**mean for IBL sections**	**mean for non-IBL sections**	**Sig.**
A. Percentage of class time with various leaders			
All student leadership	**57%**	**6%**	***
Student	28%	3%	***
Group	16%	1%	***
Class	13%	2%	**
All instructor leadership	**43%**	**95%**	***
Faculty	39%	94%	***
TA	4%	<1%	*
B. Leadership role density (number of episodes per hour of class)			
All student leadership	**4.0**	**0.4**	***
Student	2.5	0.3	***
Group	0.7	0.03	***
Class	0.8	0.06	***
All instructor leadership	**4.3**	**2.9**	*
Faculty	3.8	2.8	
TA	0.5	0.05	**
Overall variety: Total number per hour	**8.3**	**3.3**	***

*** $p<0.001$; ** $p<0.01$; * $p<0.05$

Together, these figures suggest that students played a substantial role in setting the direction and pace of an IBL class. The faculty role was less dominant but more active, as instructors did less telling and more responding, acting as a "guide on the side" rather than a "sage on the stage."

2.2.4 Students asked more questions in IBL classes.

We classified every question asked in class according to who asked the question: faculty, TA, student speaking for the first time ("new" student), or student speaking again who had previously asked a question ("repeat" student). We considered both the overall distribution of questions and the "question density," or number of questions asked per hour. In these analyses, questions asked during group work were excluded, because we could not track questions when multiple groups were working separately. In Table 2.3, Sections A-C summarize these data.

Overall, IBL classes incorporated more questions, about 50 per hour of non-group work vs. about 35 per hour in non-IBL classes. However, the distribution of who asked the questions was rather different. In both IBL and non-IBL classrooms, faculty asked the majority of all questions (65% and 79%, respectively). But in IBL classrooms, students asked more questions, especially repeat questions (12% vs. 5% of all questions asked). Comparing student participation by question density, the number of questions per hour asked by first-time speakers in IBL classes was about twice that in non-IBL classes. Repeat speakers asked nearly three times as many questions per hour as did their non-IBL peers.

Table 2.3: Question-Asking Behaviors in IBL (N=31) & Non-IBL (N=11) Classrooms

Observation variable	**mean for IBL sections**	**mean for non-IBL sections**	**Sig.**
A. Percentage of all questions asked by…			
By any student	**26%**	**19%**	
By new student	15%	14%	
By repeat student	12%	4.7%	**
By any instructor	72%	80%	
By faculty	65%	79%	**
By TA	7.7%	0.9%	**
Asker was uncategorized	1.4%	0.7%	
B. Class participation by question-asking (per class session)			
Number of students attending class	19	49	
Number who ask any questions	5.9	2.9	***
Percentage who ask any questions	**33%**	**14%**	***
C. Question density (number per hour) during all non-group work			
By any student	13	5	***
By new student	7	3	***
By repeat student	6	2	**
By any instructor	37	30	
By faculty	33	30	
By TA	4	0.5	**
Asker was uncategorized	0.6	0.2	
Total question density	**50**	**36**	
D. Percentage of all questions by type			
Recall (R)	15%	23%	*
Explanation (E)	27%	27%	
Critiquing (C)	6%	2%	**
Stretching (S)	7%	4%	
Metacognitive/checking (M)	30%	31%	
Business (B)	9%	6%	
Other (O)	2%	5%	
Type was uncategorized	5%	3%	
Highest order (C+S)	13%	6%	***
E. Question density (number per hour) by activity type			
Student presentation	42	39	
Discussion	61	52	
Computers	30		
Lecture	52	35	
Explanation	50	46	
Class business	35	37	
Other	13	26	

*** $p<0.001$; ** $p<0.01$; * $p<0.05$

TAs also had a much greater role in asking questions in IBL classrooms, asking 8% of all questions asked in IBL classrooms versus vanishingly few in non-IBL classes. This is partly because fewer of the non-IBL classes had TAs, but it also reflects the substantial teaching roles that IBL TAs took on, which we confirmed in the interview study.

Overall, IBL students were more likely to ask a question, especially a second or third question. We suggest that this is an indicator of student comfort with the classroom atmosphere: students' initial efforts to contribute or clarify were not deterred, and they felt comfortable following up or pursuing an issue.

2.2.5 More students asked questions in IBL classrooms.

In addition to asking a greater share of questions overall, IBL students also participated in greater numbers. On average, more than twice the proportion of students asked questions in an IBL class (33% of students attending) as in non-IBL classes (14% of students attending). That is, IBL students asked more questions *and* more IBL students asked questions (Table 2.3B). Again, we suggest that the fraction of students asking questions was an indicator of students' comfort in raising questions and asserting themselves to have their needs met. It may also reflect a greater level of student engagement in the mathematical topics at hand.

2.2.6 There was a modest shift toward higher-order questions in IBL classrooms.

We categorized the type of each question by its level of cognitive demand: recall (of a fact or idea), explanation (requiring a rationale or argument), critiquing (challenging an explanation, raising an issue), or stretching (extending an idea or argument to a new case). Some questions served other purposes: Metacognitive or "checking" questions concerned student thinking ("Does everyone understand?" or "Is that right?") and business questions addressed class logistics, such as "Have I received everyone's homework?" or "When is the exam?". These data are summarized in Table 2.3D-E.

There was some shift in the nature of questions asked in IBL vs. non-IBL classrooms. In IBL classrooms, the typical distribution of questions included fewer recall questions (about two thirds), but over twice as many questions of the highest orders (critiquing and stretching), compared with non-IBL classes. The overall proportion of the highest-order types was still fairly small, under 15% of all questions. We suggest that instructors found it hard to shift their questioning strategies and tended to use the same rhetorical patterns in IBL courses as they did in lecture courses. Instructors also confirmed the difficulty of "sitting on their hands" in interviews.

We expected to see a difference in the level of metacognitive questions, hypothesizing that IBL instructors would be more likely to check in on students' understanding. But there was little difference in the level of such questions, which account for about 30% in both types of classrooms. Perhaps differences in attentiveness to student understanding may be more apparent in student responses to a "checking-in" question, rather than the question itself—that is, whether students perceive the question as rhetorical or a sincere invitation to respond and seek clarification—which we did not track. The proportion of business questions was slightly higher in IBL classes, consistent with the greater time spent organizing classroom logistics.

2.2.7 In IBL classrooms, students interacted more often with each other and with instructors, and were more involved in setting the course pace and direction.

The observer survey data also indicates substantial differences in the IBL and non-IBL sections observed. These 14 indicators reflect observers' views of certain student and instructor behaviors that affect classroom interactions and atmosphere. As summarized in Table 2.4, the indicators fall into four clusters. Cronbach's alpha is given as an indication that the clusters are statistically valid and thus conceptually meaningful.

Table 2.4: Mean Observer Ratings of Classroom Interactions and Atmosphere for IBL (N=31) and Non-IBL (N=11) Sections (scale: 1=never, 5=very often)

Observation variable: Extent to which...	**mean for IBL sections**	**mean for non-IBL sections**	**Sig.**
Student-instructor interactions		Cronbach's alpha, 0.93	
Students offer ideas during class	3.54	1.93	***
Students receive personal feedback on their work	3.30	1.63	***
Instructors listen to students' ideas	4.13	2.11	***
Instructors give concrete feedback on students' work	3.44	1.58	***
Students ask questions	3.41	2.53	**
Instructors express their own ideas or solutions to problems	2.95	4.65	***
Instructors offer help to students	3.50	2.10	***
Student-student interactions		Cronbach's alpha, 0.96	
Students review or challenge others' work	3.11	1.54	***
Students work together with others	3.45	1.28	***
Students get help from others	3.37	1.47	***
Joint roles in setting course pace and direction		Cronbach's alpha, 0.69	
Students set pace or direction of class time	3.00	1.58	***
Instructors set pace or direction of class time	3.37	4.65	***
Instructor behaviors		Cronbach's alpha, 0.70	
Instructors establish a positive atmosphere	4.20	3.71	
Instructors summarize or place class work in a broader context	3.20	3.29	

*** $p<0.001$; ** $p<0.01$; * $p<0.05$

Seven items addressed the interactions of students and instructor. IBL classrooms were rated *much higher* on the extent to which students offered their own ideas during class and received personal feedback, and on the extent to which faculty in turn listened to students' ideas and gave concrete feedback on students' work. They were also rated *somewhat higher* on the extent to which students asked instructors and TAs questions, and higher on the extent to which instructors offered help. IBL classrooms were rated *much lower* on the extent to which instructors and TAs expressed their own ideas or solutions to problems. These ratings all suggest that IBL classrooms offered an environment where students could safely express ideas, ask questions, and receive feedback on their ideas.

Three items addressed student-student interactions. IBL classrooms were rated *much higher* on the extent to which students reviewed or challenged other students' work, worked together with other students, and got help from other students. These results indicate the high importance in the IBL classrooms of student-to-student conversation and collaboration, a result that is strongly supported by the student interview data.

Two items addressed students' role in setting the course pace and direction. IBL classrooms were rated *much higher* on the extent to which students set the pace or direction of class time, and *much lower* on the extent to which instructors set the pace or direction. Evidently IBL classrooms afforded students an active role in controlling the tempo and use of class time. The pace of a course was not set by an instructor's lecture delivery, but by the students' own efforts to prepare and present solutions and to comprehend the material. This did not, however, mean that students were in charge. Rather, instructors and students shared responsibility for shaping the everyday work of the course. As interview data show (Ch. 8), instructors played a critical role in setting the overarching plan for the course, selecting the curriculum, and setting the sequence of problems or tasks that students worked through.

Two items showed essentially no difference. Each reflects a behavior of instructors alone, rather than students' responses to them, and both IBL and non-IBL instructors were rated as doing these behaviors often or very often. Both types tried to establish a positive atmosphere, and both made some effort to summarize or place class work in a broader context. These high ratings, and the fact that they were so similar for IBL and non-IBL instructors alike, reflect observers' perceptions that nearly all instructors were sincere in trying to guide students through the material and doing what they could to establish a positive classroom atmosphere. That is, the differences between IBL and non-IBL classrooms were largely differences in the choice of instructional activities and **not** of instructor intent or interest in student learning.

2.2.8 Even though IBL classrooms varied widely in instructional practices, they generally differed in significant ways from non-IBL classrooms.

So far, we have compared averages of several indicators of classroom practice across a sample of IBL and non-IBL classes. Yet it is important to remember that each average represents a wide span of individual practices, and that the ranges for IBL and non-IBL classes overlapped a great deal. A few non-IBL classrooms in our sample included a good deal of student interaction; a few IBL classrooms were rather less interactive than some non-IBL sections. But as a group, the IBL classes exhibited quite distinct instructional practices from the non-IBL classes. That is, the differences between IBL and non-IBL courses were both statistically significant and meaningful.

We can thus take these differences as an evaluative measure of the project's impact as a reform effort. At the three IBL Centers for which we have observation data, instructional practices have shifted in ways that are not seen in courses not involved in the project. In the next section, we consider variability among the IBL courses studied.

2.3 Variations in Instructional Activities within IBL Course Sections

We classified sections as "IBL" or "non-IBL" based on information provided by each campus. If the course was grant-supported and considered by its home department to be "inquiry-based," we accepted that label and used it consistently in our study. The courses were designed and developed before our research team was involved. We had neither the power to design them nor any basis for criteria, prior to gathering data, to define them otherwise. In practice, however, we detected a good deal of variation, especially among IBL classes. Here we describe that variation and consider both its sources and the implications of variation.

2.3.1 The nature of instruction in IBL classes varied more widely than in non-IBL courses.

On nearly every observed variable, IBL classrooms varied more than non-IBL classrooms.

Figure 2.1: Variation in Proportions of Instructor-Centered Activities, by Section

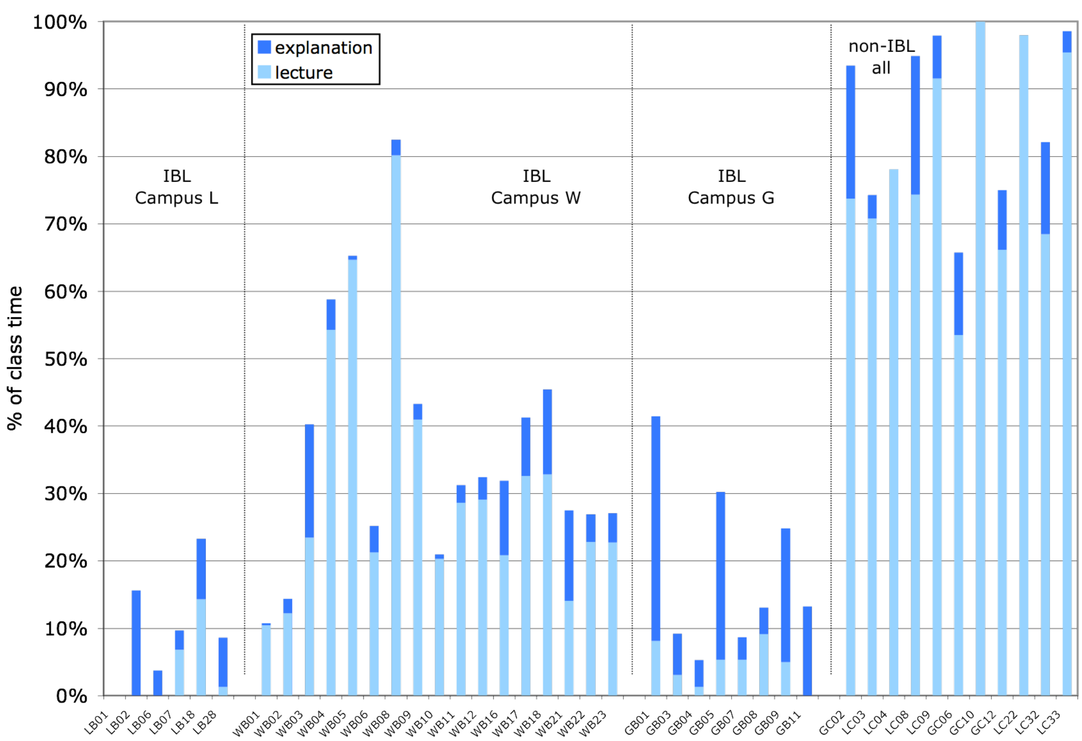
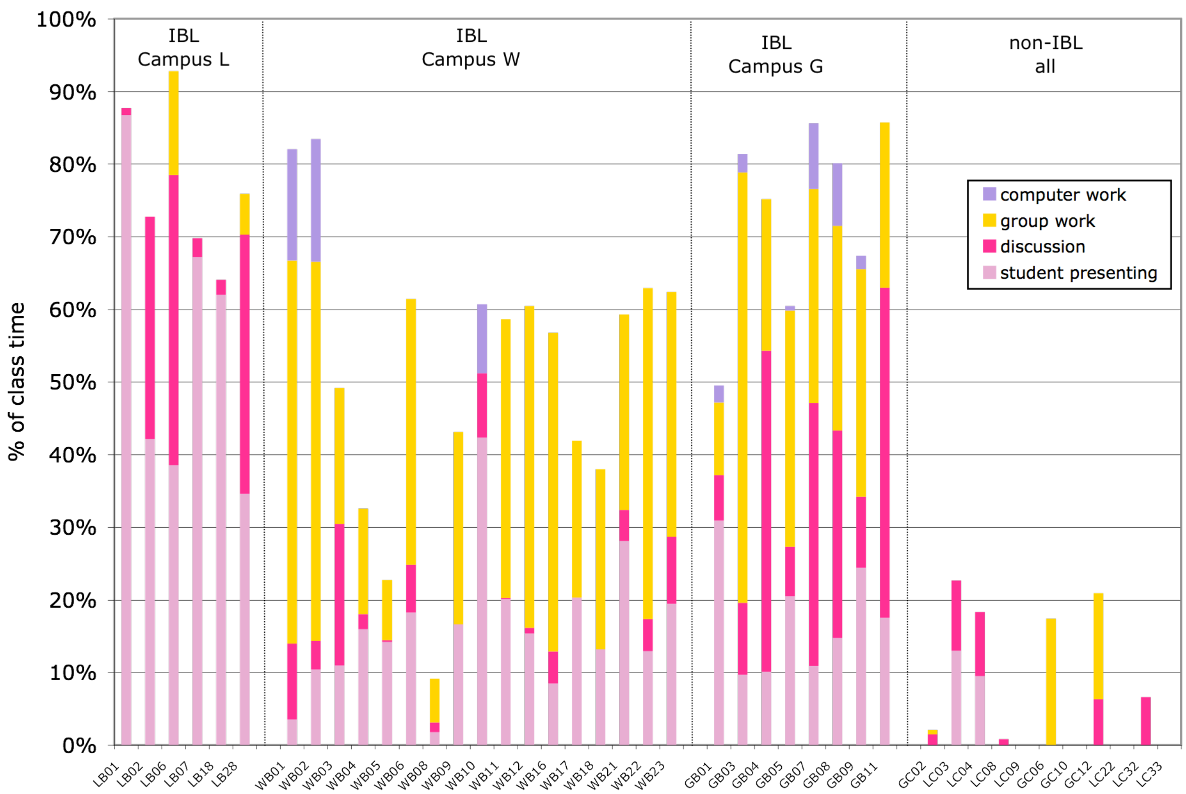


Figure 2.1 shows one example of this variation, the proportion of class time spent on instructor-centered activity across all the sections that were observed. While the figure makes plain the overall difference between IBL and non-IBL sections, it also shows the significant variation within each group. Courses varied in the total amount of instructor-centered activity, and in the mix of prepared lecture and spontaneous explanation. There are also campus-to-campus variations in the preferred style of IBL courses, as we discuss further below.

2.3.2 Multiple factors account for variation in IBL instructional practices.

The variation of classroom practice across IBL courses is a complex phenomenon with multiple sources. Figure 2.2 provides a second example of this variation, showing the proportions of class time spent on four different student-centered activities for each section studied. Like Figure 2.1 this shows both net differences between IBL and non-IBL courses, and wide variation within each. There was variation in overall student-centered uses of time, and in the mix of student-centered activities. A few non-IBL courses included some student interaction through presentations or group work, and a few courses that were designated "IBL" were nonetheless highly lecture-driven.

Figure 2.2: Variation in Proportions of Student-Centered Activities, by Section



Some of the variation was due to instructor preferences and beliefs about the best teaching approach for a given course or audience and for their own instructional style. For example, IBL classrooms tended to be either primarily presentation-focused or primarily group-work focused.

Courses that emphasized individual student presentations at the blackboard were more often aimed at upper-class students who were pursuing math, science or engineering majors. Such students may be expected to have (or to develop) the confidence to share their work with others, and the maturity to give and receive critique respectfully. Courses aimed at first-year students and at pre-service teachers more often included substantial time spent on structured group work, which can be more supportive and less exposing to students who are new to college or who are less confident about their mathematical abilities. Computer work was common in a few courses where the material was well-suited to computer simulation and problem-solving.

Variation was not just individual, but was also influenced by departmental "styles." These arose from patterns in how prior course development work was shared among faculty and how instructors new to IBL were mentored. Presentation-focused courses predominated on two campuses, while structured small group work was more common in courses at the other two.

Class size is relevant to the use of IBL methods. The mean class size for IBL courses was 19 students, versus 48 for non-IBL classes—not a statistically significant difference. The non-IBL mean of 48 masks the fact that, while a few of the non-IBL sections in this study were large lecture sections, most were similar in size to the IBL sections, with 15 to 30 students. Class size influences instructors' choices about how to teach—most IBL instructors interviewed felt that IBL worked best with 30 students or fewer—but it does not account for the observed variation.

Other variations in instruction among IBL sections probably reflect real differences in how effectively inquiry-based approaches were implemented. For example, in Figure 2.2, a few sections designated "IBL" stand out as spending rather little time on student-centered activities. Clearly there is room in these sections for a greater shift toward student-centered instruction. Combined with what we know from interviews, this shows an ongoing need for formal and informal professional development for faculty new to IBL approaches.

The overall proportion of student-centered class time is a gross indicator of the degree to which active learning can be taking place, but the observation data include also more subtle indicators of the effectiveness of implementation. Within courses that dedicated substantial class time to student-centered instructional activities, whole-class discussion appeared to be distinctive. The proportion of class time spent in discussion was quite variable, from zero to over 40%. Yet the presence or absence of discussion was not clearly linked to campus or audience patterns in the same manner as was, say, group work.

From a detailed look at classes where the time spent in whole-class discussion is substantial, we saw that activities tended to regularly alternate in a way that suggests discussion "burst out" in response to structured small group work or a student presentation. Group work, lecture, and student presentations were pre-planned activities, but discussion seemed to be more spontaneous. Indeed, the sample includes examples of discussion-rich sections that emphasized each of the three main instructional activities: group work, student presentation, or lecture. Discussion thus may be a somewhat independent indicator of an interactive classroom. It may reflect a particularly open atmosphere or relate to an instructor's skill in spotting "teachable moments" that can be used to spark student conversation.

In Chapter 4, we will discuss the linkage between variability in classroom practice, as measured by several observation variables, and student outcomes, as measured by student learning gains data from surveys. That analysis draws on both this chapter and the survey data of Chapter 3.

2.4 What is Inquiry-Based Learning? Patterns of Practice in this Project

The range of variation in classroom practice across IBL sections is noteworthy but predictable. Teaching and learning are human activities that are inherently rich and messy. We are often asked to "define" IBL—but in fact the definition that matters was provided by the instructors who taught these courses. Instead, we can characterize the range of practices. Within this project, we identify certain key features by both their statistical frequency and their stated importance to students and instructors:

- The main work of the course was problem-solving: students solved challenging problems alone or in groups, in and out of class. In class they shared, evaluated and refined their own and each others' solutions.
- The course was driven by a carefully built sequence of problems or proofs, rather than a textbook. The pace of the course was set by students' movement through this sequence, rather than pegged to a pre-set schedule.
- Course goals tended to emphasize development of skills such as problem-solving, communication, and mathematical habits of mind, as much or more than covering specific content.
- To accomplish these goals, most of class time was spent on student-centered instructional activities. Students or groups of students played a leading role in guiding these activities. Most activities were conducted for just a few minutes at a time: class work tended to change gears often and switched frequently between activities.
- Instructors' main role was not to lecture. They might give mini-lectures to shape or sum up the day's work, set up a group activity, or provide context for a set of theorems. Instructors (as well as other students) might offer impromptu explanations to respond to a comment or question. They might ask questions to clarify student thinking, refine a presented solution, give feedback, or to elicit such comments from other students.
- Student voices were heard in the classroom: presenting, explaining, arguing, asking questions. Their active participation meant that students themselves had much control over how class time was spent and how fast the class moved through the material.
- This joint responsibility for the depth and progress of the course fostered a collegial atmosphere that placed value on respectful listening and critique and invited every class member to contribute fruitfully to the mutual development of mathematical ideas.

What is not visible in the observation data—which were gathered only during the conduct of class—is the behind-the-scenes structure of an IBL course. We do know something about this from interviews with instructors and students. Outside class, much of students' work time was spent preparing for these activities: solving problems or deriving proofs to present or discuss in groups. In interviews, students reported spending many hours on this work. Instructors invested substantial time in constructing the "script" or sequence of problems or (if they had received a script from another instructor) in understanding it and tweaking it. When a script was available it was much less work to teach the course than if they developed it themselves. Both faculty and TAs reported that they held many office hours outside class, and that students made use of them. Students likewise reported that the availability of timely help was important to making progress.

2.5 Conclusion: Strengths, Limitations, and Future Directions

These analyses set the scene for our research results on student learning: We must show that students in IBL classes experienced something different even to bother studying the impact of these courses on their learning. We find that most IBL sections did offer a quite different set of instructional activities, and that the classroom atmosphere and interactions were also rather different. Even though there were variations among the courses designated "IBL," the net differences were large. In later chapters we will discuss the impact of all these observed classroom practices on students' learning and instructors' teaching, link them to other data, and further elucidate how these processes help or hinder student learning.

There are some limitations to the study as well, especially in the size and nature of our samples. The non-IBL sample is small, because comparative non-IBL sections were available only on two of the three campuses (and not for every IBL course) and because some non-IBL instructors chose not to participate. The fourth campus did not participate in this study, and its practices are thus not at all represented. While the observation protocol was deliberately simple to allow for multiple observers and to keep costs low, more refined data could be obtained from a more detailed protocol, stronger observer training, and more effort to ensure inter-rater reliability. Observers found it difficult to classify question types on the fly, and these data appear (so far) to provide fewer insights into the nature of IBL instruction than do other observation data.

These analyses also raise many questions for future study. In future analyses and writing beyond this report, we plan to examine these data in greater depth, especially the relationship between different observation variables to each other and to an important source of data that we have not discussed here, the observers' written notes about the classes. We will also analyze the utility of this observation protocol for future studies and make recommendations for its improvement. Lastly, we hope to link these data firmly to student outcomes. While Chapter 4 reports on some preliminary work to clarify this linkage, more work remains to be done to fully understand how student learning outcomes depend on their classroom experiences.

Chapter 3: Findings from Student Surveys on Learning Gains, Course Experiences, and Attitudes

3.1 Introduction

Using survey methods, we could efficiently gather a lot of information about student outcomes at low cost. Two different survey instruments were used to measure student outcomes and to compare these outcomes between various student groups. The attitudinal survey was designed to probe the nature of students' mathematical beliefs, affect, learning goals, and problem-solving strategies, and changes in these as a result of an IBL course. The learning gains survey or SALG-M (SALG, 2008) examined students' experiences of class activities and their cognitive, affective and social gains from a mathematics class. The surveys addressed the following research questions:

- What learning gains do students report from an IBL mathematics class?
- How do students experience IBL class activities? How do students' class experiences account for their gains?
- What kind of beliefs, affect, goals and strategies do IBL students report at the start of a mathematics course?
- How do these approaches change during a college mathematics course? How do these changes relate to or explain students' learning gains?
- For each of these outcomes—learning gains, experiences, attitudinal measures, and changes—how do the outcomes for IBL students differ from those of non-IBL students, and among IBL student sub-groups?

The survey instruments, samples, and analysis are detailed in Appendix A3. These analyses are based on data from 57 IBL and 18 non-IBL sections on all four campuses.

The SALG-M was given only at the end of a course, when students can describe their experiences during a math class and what learning gains they made from the course. Survey questions include both multiple-choice and open-ended questions. The learning gains analyses include data from 563 IBL math-track students, 288 non-IBL math-track students, 220 IBL pre-service teachers, and 34 non-IBL pre-service teachers.

The attitudinal survey was a pre/post instrument, where students answered the same questions at the start and end of their course. Pre-survey data show students' beliefs, motivations, and earning strategies at the start of their course. Comparison with the attitudinal post-survey data reveals changes in students' attitudes. Analyses of the attitudinal pre-/post-survey data are based on information from 754 matched surveys: 390 IBL math-track students (defined in 1.6.1), 156 non-IBL math-track students, 184 IBL pre-service teachers, and 24 non-IBL pre-service teachers.

In some cases, we break out results for different student subgroups, e.g. by gender or year in school. For math-track students, we can often do this, but for pre-service teachers, the number of respondents is often too small to present such analyses. For example, there are few male pre-service teachers in the sample, so we cannot analyze this group by gender. Moreover, as a group, IBL and non-IBL pre-service teachers cannot be directly compared. The IBL group comes mainly from courses specifically targeted to teacher preparation, while the non-IBL group represents math majors pursuing high school teaching and taking the same advanced courses as other math majors. Their course experiences and outcomes are thus not comparable.

Below we report results from the SALG-M instrument on student learning gains, followed by results on student course experiences. We then discuss results from the attitudinal survey: first students' initial attitudes and beliefs, then changes in these.

3.2 Self-Reported Learning Gains

Analysis of students' self-reported learning gains provided information about the effects of IBL instructional practices on their cognitive, affective and social learning of mathematics. In addition, we compared IBL students' learning gains to non-IBL students' gains. First, we report results from the structured survey questions and about findings based on open-ended questions.

Throughout this report, we discuss three broad types of learning gains:

- cognitive gains: growth in understanding mathematical concepts and ideas, in thinking and problem-solving, and in applying knowledge outside classroom
- affective gains: enhanced confidence, increased positive attitude toward mathematics, and persistence in studying mathematics and solving problems
- social gains: gains in collaboration with other students and comfort in teaching.

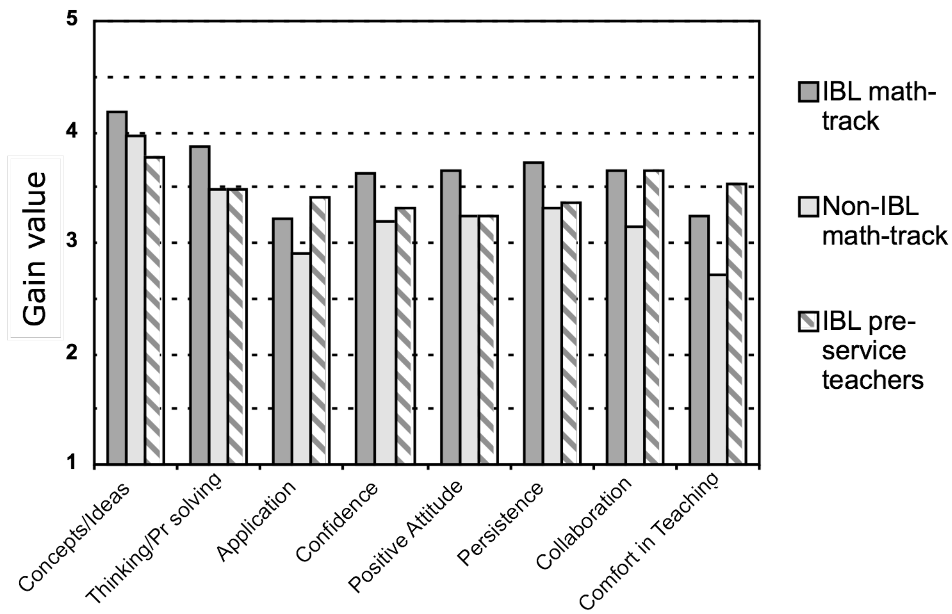
Section 3.2.1 discusses quantitative items measuring these gains, highlighting differences that are statistically significant. Section 3.2.2 describes some additional gains based on students' spontaneous comments about changes in their ways of learning mathematics, communication skills and enjoyment of mathematics. In later sections, we discuss differences in these gains reported by different student sub-groups.

3.2.1 IBL students reported higher cognitive, affective and social gains.

Figure 3.1 displays average ratings of learning gains for three main student groups: IBL math-track students, non-IBL math-track students, and IBL pre-service teachers. It shows that learning gains were rather high for all the students. However, there were statistically significant differences in the averages among these three student groups for all three types of learning gains. In every case, these differences favor IBL students.

IBL math-track students reported the highest gains in understanding mathematical concepts or ideas ($p < 0.001$) and in mathematical thinking and problem solving ($p< 0.001$). Their gains in applying mathematical knowledge also exceeded those of non-IBL students ($p< 0.001$). That is, math-track students' outcomes in learning mathematical concepts and in thinking are better in IBL sections than in non-IBL sections. The effect of IBL instruction on cognitive gains was less strong for pre-service teachers, but their gains in applying mathematics were the highest. IBL classes for this group typically focused more on using mathematical ideas in other contexts, especially in teaching school mathematics.

Figure 3.1: Average Learning Gains by Student Group



Among affective learning outcomes, IBL math-track students reported the highest gains in confidence ($p< 0.001$). They also reported higher gains than other students in positive attitude towards mathematics ($p< 0.001$). Pre-service teachers reported lower gains in these areas, only slightly higher than non-IBL students and lower than IBL math-track students ($p< 0.001$). This likely reflects their general lower confidence and more negative attitudes to begin with, as is commonly reported in the literature. Similar group differences applied to gains in persistence, where IBL math-track students reported higher gains than the other students ($p< 0.001$).

That is, among math-track students, both cognitive and affective learning outcomes were higher in IBL sections than in comparative non-IBL sections. For pre-service teachers, cognitive and affective learning outcomes were less strong.

Collaboration was a significant feature in IBL instruction, and both IBL math-track students and IBL pre-service teachers reported clearly higher gains in working with others than did math-track non-IBL students ($p< 0.001$). Their gains in comfort in teaching mathematics also clearly exceeded those of non-IBL math-track students ($p<0.001$). Thus IBL class activities increased students' skills in cooperating with and teaching other students in ways that did not occur in lecture-based courses. Students became more willing and able to discuss mathematics and present mathematical ideas to other people. The results are strengthened by our findings from student interviews.

3.2.2 IBL students reported more gains in their spontaneous comments.

The findings from numerical survey items are corroborated by students' spontaneous written comments in response to several optional, open-ended questions on the SALG-M. Students described their gains in their conceptions of mathematics, in learning mathematics and solving mathematical problems, and in communication. Table 3.1 shows the number of write-in comments, sorted by category and by student group.

Table 3.1: Learning Gains Reported in Written Responses, by Student Group: Number of Written Comments and Percentage of Students Responding

Type of gain reported in written comments	IBL math-track students	Non-IBL math track students	IBL pre-service-teachers	Sig. (freq. distribution)
Cognitive Gains	*208 (27%)*	*88 (15%)*	*67 (25%)*	***
Learning & understanding math concepts	145 (19%)	27 (9%)	47 (17%)	***
Thinking and problem-solving skills	101 (13%)	40 (7%)	22 (8%)	**
Changes in conceptions of mathematics	40 (5%)	35 (6%)	12 (4%)	
Changes in Learning	*166 (21%)*	*55 (9%)*	*77 (29%)*	***
Changes in ways of learning math & solving problems	72 (9%)	31 (5%)	58 (22%)	***
Independence in learning & solving problems	73 (9%)	16 (3%)	11 (2%)	***
Learning from others	41 (5%)	10 (2%)	24 (9%)	***
Affective Gains	*59 (8%)*	*10 (2%)*	*16 (6%)*	***
Confidence in own math ability & skills	27 (3%)	5 (0.9%)	12 (4%)	**
Enjoyment of mathematics or learning mathematics	20 (3%)	2 (0.3%)	5 (2%)	*
Communication skills	67 (9%)	10 (2%)	30 (11%)	***
Percentage of students reporting any write-in gain	**36.8%**	**21.4%**	**41.6%**	
Percentage of students reporting 3 or more write-in gains	**19.1%**	**7.7%**	**20.1%**	***

** $p < .05$*, $p < .01$**, $p < .001$****

Overall, 13 to 38% of IBL math-track students and 12 to 20% of non-IBL math-track students chose to write in a comment about their learning gains (depending on the question). Compared to their non-IBL peers, more IBL students wrote in any gain. In all, only about 20% of non-IBL math-track students wrote any comment, but about 40% of both IBL groups wrote in a comment. IBL students also wrote in more gains than non-IBL students. About 20% of both IBL groups wrote at least three distinct comments on their gains, but only 8% of non-IBL math-track students did so (p< 0.001).

As Table 3.1 displays, statistically significant differences among the three student groups were found in the number of write-in gains reported in nearly every category. In particular, IBL math-track students reported one or more gains across several categories more often than did non-IBL math-track students.

Among cognitive gains, IBL (19%) more than non-IBL math-track students (9%) reported better knowledge and understanding of mathematical concepts. For example, they (8%) more often than non-IBL (1%) students wrote about better recall of mathematical knowledge. "I remember [concepts] much better because I come up with them on my own," wrote one IBL student. Another wrote, "I had to work on the key ideas instead of them being told to me; therefore I understand and remember these ideas very well."

Enhanced mathematical thinking and problem-solving skills also were mentioned more often by IBL than non-IBL students. For example, an IBL student reported his/her main gains as "the rigorous reason and the approaches to conceptualizing problems." Another IBL student carried out "the logical process necessary to do these proofs." Gains in general thinking skills were reported in statements like one student's description of learning "how to think of things in multiple perspectives." Relatively few IBL and non-IBL students wrote about the usefulness of their math course for other courses or academic areas.

Among learning changes, IBL pre-service teachers emphasized changes in their ways of learning mathematics and problem-solving, and IBL math-track students commented on their gains in independence. IBL students also more often described their appreciation of learning from other students and their gains in communicating about mathematics.

Overall, students commented less about their affective gains than other types of gains, but IBL students wrote slightly more often about their increased confidence in and enjoyment of mathematics. No difference was found in students' reports of the ways that they understand the nature of mathematics.

These results confirm the differences in learning that were observed in the structured survey responses. Generally, IBL students reported higher cognitive, affective, and social gains as compared to non-IBL students. They also described changes in the ways of mathematics learning and problem-solving more than did non-IBL students. Clearly, IBL instructional practices had a powerful and positive effect on students' learning of mathematics.

3.2.3 Women strongly benefited from IBL classes in comparison with traditional approaches.

We compared learning gains for women and men in math-track classes. Both IBL men and women reported clearly higher gains than did non-IBL men and women. Moreover, among IBL students, women's cognitive, affective and social gains were on average as high or higher than those of men, while their gains in collaboration exceeded that of IBL men ($p < 0.01$). No gender difference appeared in gains in teaching.

However, the situation in non-IBL courses is quite different. Non-IBL women reported lower gains than non-IBL men in several areas: understanding mathematical concepts ($p < 0.05$), mathematical thinking and problem-solving ($p < 0.01$), confidence ($p < 0.05$), and positive attitude toward mathematics ($p < 0.01$). It appears that traditional methods of teaching college mathematics prevented women from learning and developing confidence and positive attitudes about mathematics. This is an important result: the use of active IBL instructional methods helped both men and women but particularly benefited women's study of college mathematics.

The important role of IBL instruction in diminishing gender differences in mathematics is also evident when comparing IBL women's and non-IBL women's learning gains. Table 3.2 displays average learning gains for IBL and non-IBL women, and for IBL and non-IBL men, including results of statistical tests for group differences.

Table 3.2: Average Learning Gains for Math-Track Students, by Gender

Type of learning gain		AVERAGE			AVERAGE		
		W1 IBL women	**W2** Non-IBL women	Sig.	**M1** IBL men	**M2** Non-IBL men	Sig.
Cognitive	*Math concepts*	4.3	3.8	***	4.2	4.0	
	Math thinking	3.9	3.2	***	3.9	3.6	**
	Application	3.2	2.8	*	3.2	3.0	**
Affective	*Confidence*	3.7	3.0	***	3.6	3.3	***
	Positive attitude	3.7	2.9	***	3.6	3.3	**
	Persistence	3.8	3.2	***	3.7	3.4	**
Social	*Collaboration*	3.9	3.2	***	3.5	3.2	***
	Teaching	3.3	2.7	**	3.2	2.7	***

** p < .05*, p < .01**, p < .001****

As shown in Table 3.2, IBL women reported clearly higher gains than did non-IBL women. The largest differences lie in cognitive gains in understanding mathematical concepts and in mathematical thinking, affective gains in confidence and in positive attitude towards mathematics, and in collaboration. The differences between IBL and non-IBL men are similar but somewhat weaker than among women.

Importantly, the difference between IBL and non-IBL women's gains did not depend on their prior mathematics achievement. IBL women reported higher cognitive, affective, and social gains than did non-IBL women even after controlling for their prior GPA and their expected grade in the start of the course ($p < 0.01$). In contrast, the differences between IBL and non-IBL men's gains were related to IBL men's prior GPA level or higher expected grade as compared to that of non-IBL men. These findings apply to smaller data sets of mostly older students, since first-year students did not report a GPA. However, analysis with the larger data set based only on post-survey results and demographics, indicated that IBL women reported higher gains than did non-IBL women after controlling for their expected grade at the end of their course ($p < 0.001$).

The latter result also, but less clearly ($p< 0.05$), applied to math-track men. These results indicate a real positive effect of IBL classes on women's learning of college mathematics that is independent of their prior academic background.

Gender differences were also seen in spontaneous write-in comments. IBL women wrote more (22%) about gains in understanding mathematical concepts than did non-IBL women (5%; $p< 0.001$). Moreover, unlike non-IBL women, IBL women described gains in confidence. Likewise, IBL men's written comments also showed higher gains. They (14%) more often than non-IBL men (7%) reported gains in mathematical thinking ($p< 0.05$), independence ($p< 0.01$), learning from other students ($p< 0.01$), and enjoyment of mathematics ($p< 0.05$). Again, these spontaneous observations confirm that IBL instruction has a positive influence on both women' and men's learning of college mathematics.

3.2.4 Students early in their college careers benefited more from IBL instruction than did later-stage students.

To see whether the benefits of IBL instruction depended on students' level of college experience, we divided math-track students into three roughly equal groups: first-year students, a mid-career group of sophomores and juniors, and a late-stage group of seniors and graduate students. Table 3.3 shows average learning gains for these three groups of different academic status.

Table 3.3: Average Learning Gains for Math-Track Students, by Academic Status

Type of Gain		First-year students			Mid-career students (sophomore, junior)			Late-career students (senior, graduate)		
		IBL	Non-IBL	Sig.	IBL	Non-IBL	Sig.	IBL	Non-IBL	Sig.
Cognitive	*Concepts*	4.2	4.0	*	4.3	3.8	***	4.1	4.0	
	Thinking	4.0	3.5	***	3.9	3.4	**	3.7	3.3	
	Application	3.3	2.9	***	3.2	3.0		3.0	2.7	
Affective	*Confidence*	3.6	3.3	**	3.7	3.2	***	3.5	2.7	*
	Positive attitude	3.7	3.3	***	3.8	3.2	***	3.5	2.9	*
	Persistence	3.8	3.5	***	3.7	3.2	***	3.5	2.9	
Social	*Collaboration*	3.9	3.3	***	3.6	3.0	***	3.4	2.7	**
	Teaching	3.2	2.8	**	3.3	2.6	***	3.1	2.7	

** p < .05*, p < .01**, p < .001****

First-year IBL students gained more than did late-stage IBL students in mathematical thinking ($p< 0.05$) and persistence ($p< 0.05$). They also reported higher gains in collaboration as compared to the mid-career ($p< 0.05$) and late-stage groups ($p< 0.01$). In addition, relative to older students, they reported greater gains in mathematical thinking and in the ability to persist in solving problems. For first-year students, such gains in thinking and problem-solving, and in their ability and inclination to work with others, may be especially important if they carry over to their study of mathematics during later college years.

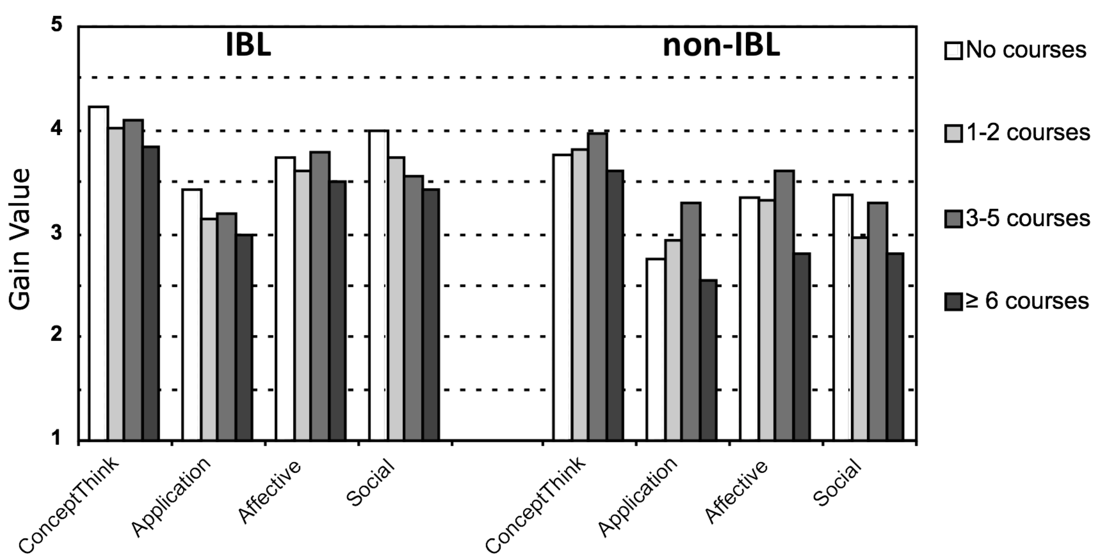
Differences in learning gains between IBL and non-IBL math-track students were most clear among first-year students and the mid-career group (see Table 3.3). IBL first-years clearly gained more than their non-IBL peers in several areas: mathematical thinking, application of mathematical knowledge, positive attitude, persistence in studying mathematics, and working with others. These differences between IBL and non-IBL were also valid in the mid-career group: IBL sophomore/juniors' gains in understanding mathematical concepts and teaching mathematics exceeded those of non-IBL sophomores and juniors.

Finally, although the oldest IBL students also gained more than their non-IBL peers, the differences were not as large. Among late-stage students, IBL students' affective gains ($p < 0.01$) and gains in collaboration ($p < 0.05$) exceeded those of non-IBL students. While younger students benefited most from IBL instruction, it had positive effects on students' learning at all stages.

3.2.5 Students who had previously taken fewer mathematics courses benefited more from IBL instruction than did experienced students.

A different analysis also examined the impact of IBL depending on students' prior experience, comparing learning gains according to the number of prior college mathematics courses taken. Here, math-track students were divided into four groups: those with no previous math courses, 1-2, 3-5, and 6 or more previous courses. Figure 3.2 displays averages of learning gains for these groups, separately for IBL and non-IBL math-track students. The findings confirmed the idea of higher gains for younger IBL students.

Figure 3.2: Average Learning Gains by Number of Prior College Mathematics Courses, for Math-Track Students



Unlike non-IBL students, IBL math-track students who had no previous college mathematics course reported higher cognitive and social gains than did more experienced IBL students. Table 3.4 summarizes the statistically significant differences. The least experienced students made broad gains across all types of gains. More experienced students made some gains in attitudes and social skills that may support their mathematics learning and transfer to other areas. It is also notable we have no evidence that IBL measures do harm.

Table 3.4: Summary of Differences in IBL vs. Non-IBL Student Outcomes, by Prior Course Experience

Type of gain	No prior college math courses	1-2 prior college math courses	3-5 prior college math courses	6 or more prior college math courses
cognitive gains	+ ($p < 0.01$)			
affective gains	+ ($p < 0.05$)			+ ($p < 0.01$)
social gains	+ ($p < 0.001$)	+ ($p < 0.01$)		+ ($p < 0.05$)

While students at all levels benefit more from IBL instructional approaches than from traditional teaching, the benefits to younger and inexperienced students are the strongest. Nonetheless, the most advanced and experienced students benefited from IBL instruction in the areas of collaboration, confidence, and persistence.

3.2.6 Lower-achieving students benefited more from IBL classes in comparison with traditional instruction.

We also wondered whether students' learning gains varied by their prior achievement level, measured using the overall college GPA they reported on the pre-survey. Correlations between self-reported learning gains and overall GPA were weak, but suggested that the relationship between students' prior academic record and their learning gains in these courses differed between the IBL and non-IBL groups. For further analysis, we divided the students into three groups: low-achieving (GPA below 3.0), medium-achieving (GPA 3.0 to 3.79), and high-achieving (GPA 3.8 or above),[1] and compared these groups for IBL, non-IBL math-track students, and IBL pre-service teachers.

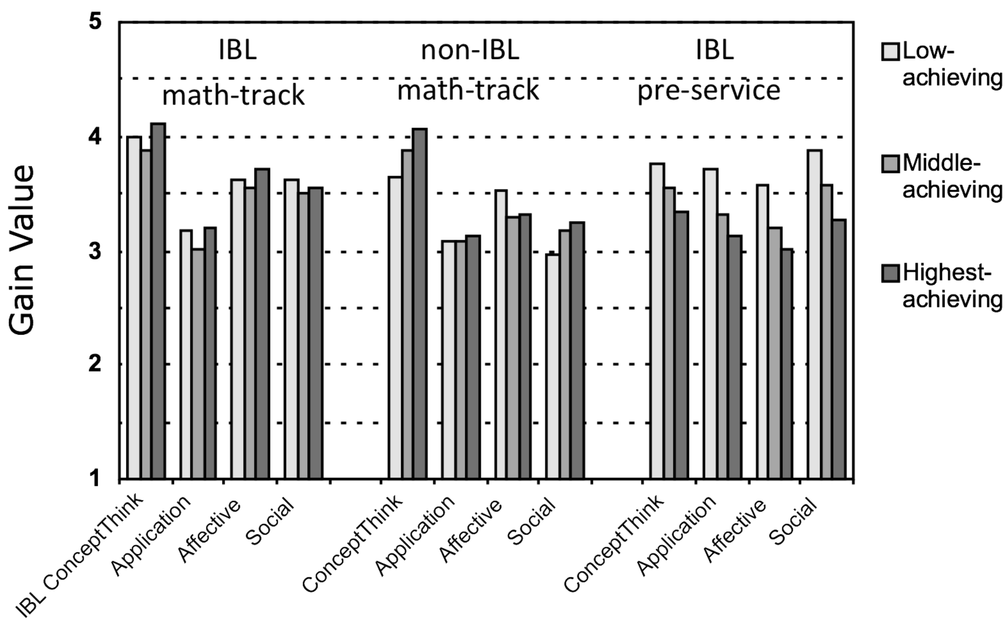
Figure 3.3 shows that highest and middle-achieving non-IBL math-track students reported higher gains than did low-achieving non-IBL students in mathematical concepts and thinking, and in collaboration. That is, traditional teaching tended to benefit the students who were already strong. In contrast, low-achieving IBL math-track students reported slightly higher cognitive, affective, and social gains as compared to IBL middle-achievers. That is, lower-achieving students had a small boost. However, the number of students in each subgroup was rather small and these differences were not statistically significant.

The averages in Figure 3.3 indicate that the benefit of IBL instruction for low-achievers was more apparent among IBL pre-service teachers. All the learning gains were highest for low-

[1] Students reported their GPA by selecting from multiple categories. The low-achieving group includes those who marked 'below 2.0', '2.0 to 2.49', or '2.5 to 2.99'. The middle-achieving group marked the choices '3.0 to 3.49' or '3.50 to 3.79'. The high-achieving group marked the answer '3.8 or higher.'

achievers. Moreover, middle-achieving IBL pre-service teachers reported higher gains than their highest achieving IBL peers. This trend was reflected also in the slight negative correlations between pre-service teachers' learning gains and their overall GPA. Even though the differences were not statistically significant, these results are worth noting since K-12 pre-service teachers' low mathematical skills often impede their own teaching of mathematics (Capraro et al., 2005; Hill, Rowan, & Ball, 2005; Ma, 1999; Wilkins, 2008). IBL experiences may enhance the mathematical understanding needed for teaching, as well as improve future teachers' attitudes toward mathematics, confidence in their own mathematical abilities, and social learning skills. Further studies on larger, comparative samples would be worthwhile.

Figure 3.3: Average Learning Gains, by GPA Level and Student Group



We also compared IBL and non-IBL students' gains within each achievement level. Among math-track students, low-achieving IBL students gained slightly more than their non-IBL peers in three areas—confidence, collaboration, and mathematical concepts and thinking—but only the difference in collaboration gains was statistically significant ($p< 0.05$). IBL approaches appear to improve lower-achieving students' mathematics learning in particular, while traditional methods benefit high-achieving students modestly more than lower-achieving students.

3.2.7 Students' race, ethnicity and academic major did not affect their learning gains.

We also looked for variation in IBL students' learning gains with respect to other demographic characteristics, including their ethnicity, race, and academic major.

First, we compared students of Hispanic and non-Hispanic ethnicity. This analysis has low statistical power because the number of Hispanic students was very low. For IBL math-track students, no statistically significant difference was found in the learning gains of Hispanic (30) and non-Hispanic IBL (342) students. Likewise, for non-IBL math-track students, there was no difference for Hispanic (14) or non-Hispanic (136) students. However, among IBL pre-service teachers, Hispanic students' (16) gains in collaboration exceeded those of their non-Hispanic peers (169; $p < 0.05$). Collaborative IBL instructional methods may be especially beneficial for Hispanic students, but this result is based on a small sample size.

We also examined differences by race. We could only compare white and Asian students, because the number of other students of color in our sample was very low (see Appendix A3). Asian IBL students' gains were slightly higher than other IBL students' gains, but the differences were not statistically significant. Asian IBL math-track students' gains were also slightly higher than that of their non-IBL peers, and these differences were statistically significant for gains in two areas, mathematical thinking and confidence ($p < 0.05$). In contrast, non-IBL Asian math-track students gained on average at the same level as their white non-IBL peers (especially in mathematical concepts and thinking). These findings suggest that IBL teaching approaches do not solely benefit white students; again, further investigation on larger, more racially diverse samples would be helpful.

Finally, we considered differences by academic major. Most math-track students were pursuing degrees in pure or applied mathematics. Among IBL math-track students, learning gains did not vary by major: math, science, engineering/computer science (CS), economics, or non-science.

We also compared the gains of IBL vs. non-IBL students by major. Interestingly, among science majors, there was no difference between IBL and non-IBL students' gains, but engineering/CS majors gained more than their non-IBL peers in confidence ($p < 0.05$). IBL economics majors reported higher cognitive, affective, and social gains as compared to their non-IBL peers ($p < 0.05$), but the number of economics majors was very small. Benefits from IBL college mathematics teaching may vary modestly depending students' academic interests.

3.3 Students' Experiences of Instructional Activities and Practices

3.3.1 Active participation and interactions in class helped student learning in IBL classes.

As our observation data show (Chapter 2), in-class activities varied a lot between lecture-oriented, traditional teaching and active IBL instructional methods. Students' out-of-class study activities also differed. IBL and non-IBL students thus experienced their college mathematics class differently. Important linkages between classroom observation data and students' learning gains are presented in Chapter 4. Here we report students' responses to survey items asking about the impact on their learning of particular instructional practices, including the overall instructional approach, classroom activities, tests or other assignments, and interactions during the class. These question were framed as "How helpful was… to your learning?". Students' ratings of these practices are listed in Table 3.5.

Both IBL and non-IBL students generally found the instructional practices in their mathematics class helpful to their learning, but their indications of which practices were particularly helpful are noteworthy. IBL students reported the greatest help to their learning from active participation and interaction during class work. In contrast, non-IBL math-track students rated these activities the least helpful. IBL students also considered interactions with their instructor, TA, and peers to be very helpful. Students clearly benefited from their own active involvement and interactions during IBL class work.

Table 3.5: Average Ratings of the Helpfulness of Classroom Practices, by Student Group

Classroom Experience Variable		Average Rating				
		math-track students		**pre-service teachers**		Statistically significant group differences
		A IBL	**B** Non-IBL	**C** IBL	**D** Non-IBL	
Classroom practices (1=no help to 5=great help)	*overall approach*	3.8	3.6	3.6	3.7	A > B ***, C *
	active participation	3.9	3.2	3.9	3.8	A***, C ***, D ** > B
	individual work	4.0	3.9	3.6	3.9	A***, B *** > C
	tests	3.3	3.3	3.0	3.3	A**, B ** > C
	other assignments	3.7	3.4	3.3	3.4	A > B***, C **
	interactions	3.9	3.5	3.9	3.7	A***, C *** > B
	- with instructor	3.9	3.3	3.5	3.5	A > B***, C ***
	- with TA	3.8	3.5	3.8	3.8	A **, C * > B
	- with peers	4.0	3.6	4.2	4.2	A***, C ***, D ** > B

** p < .05*, p < .01**, p < .001****

Students' views of the helpfulness of different types of course work also varied. IBL students reported a lack of traditional tests in their classes more often than did non-IBL students (16.6% vs. 0.3%)--consistent with instructors' reports of some degree of shift in assessment practices (Ch. 8). Math-track IBL students rated tests as highly as did students in non-IBL sections but found other assignments clearly more helpful than tests. Moreover, they considered other assignments and homework more helpful than did math-track non-IBL students. Compared with the other student groups, IBL pre-service teachers rated tests and individual work as less helpful. This result is consistent with their stronger preference for and beliefs about group work, as seen in other survey items (see Section 3.4.1). However, all students considered individual work to be very helpful for their learning, reflecting the general importance of students' own effort and thinking for successful mathematics learning.

3.3.2 Active teaching and learning practices contributed to IBL students' learning gains.

To understand how IBL students' gains related to their course experiences, we used regression analysis. Here we use the statistical measure "variance" as a measure of how much the student gains are explained or predicted by the different variables describing classroom practice.

Among IBL math-track students, two types of gain, affective gains and gains in mathematical concepts and thinking, were best explained by student ratings of "other assignments," which accounted for 35% and 38% of the variance, respectively. Four other variables—active participation, interaction with instructors, tests, and individual work—explained an additional 10% of the variance in students' gains in mathematical concepts and thinking and an additional 9% of the variance in their affective gains.

Interaction with peers (34%) explained the largest portion of math-track students' gains in collaboration. Student responses about other assignments, interaction with instructor, and active participation explained an additional 9% of the variance in collaboration. These results suggest that the assignments used in IBL classes clearly contributed to student learning. In addition, both interactions with instructor and peers and students' individual work on problems helped their learning of mathematics and boosted their confidence, attitude and persistence. These results are consistent with—and enriched by—student descriptions of the important roles of collaborative and individual work in their learning processes (Ch. 7).

The results for IBL pre-service teachers differed somewhat. Their gains in mathematical concepts and thinking were best explained (38%) by active participation "helping me learn." But their affective (51%) gains were best explained by how helpful they found tests. Individual work and interaction with instructors were other significant contributors to these gains. Active participation best explained (39%) student gains in collaboration, together with additional (12%) effects due to interaction with peers, other assignments, and tests. Overall, pre-service teachers reported that their own active participation in particular helped their learning of mathematics and enhanced their work with other students. Unlike among IBL math-track students, the nature of tests had a significant role in the development of pre-service teachers' confidence, attitude and persistence in mathematics. These results also highlight how the parts of an IBL course worked together to secure student learning gains.

Among non-IBL math-track students, learning gains were less well explained by the course experiences that this survey probed. Their gains in mathematical concepts and thinking (39%) and affective gains (37%) were mostly explained by the effect of tests. Interactions with instructor and peers explained an additional 9% of the variance in their cognitive gains and interactions with peers an additional 2% in their affective gains. Their gains in collaboration were best explained by their experience of tests (23%), and additionally (13%) by their interaction with peers and their own active participation. These results may indicate that learning was less focused on everyday classroom experiences, and more often organized around the focusing event of preparing for tests.

Together, these results suggest that tests played a more important role in non-IBL students' class experiences and resulting learning gains than they did for IBL students. Tests also influenced IBL pre-service teachers' learning gains. In contrast, for IBL math-track students, everyday learning experiences such as class assignments besides tests better accounted for learning gains. IBL students also found active participation and peer interactions very helpful for their learning.

3.3.3 Math-track men and women had more positive experiences in IBL sections.

For math-track students, we compared IBL students' course experiences with those reported by non-IBL students, broken out by gender. Both men and women in IBL sections rated the overall instructional approach more helpful than did their peers in non-IBL sections. This difference was even clearer among women ($p < 0.001$). Both men and women rated class interaction and their own active work as more helpful than did non-IBL men and women ($p < 0.001$).

Similar differences in favor of IBL sections applied to students' experiences of assignments other than tests. IBL men ($p < 0.05$) and especially IBL women ($p < 0.001$) rated these assignments as more helpful than non-IBL students. These results may help to explain our observation of generally stronger learning gains for women. IBL course activities helped both women and men to learn mathematics, but these contributed even more to women's learning.

3.3.4 Younger students had more positive experiences of active participation and interaction.

We compared students' course experiences between first-year, mid-career (sophomore/junior), and late-stage (senior and older) students. Overall, IBL first-year students had more positive experiences of their own active participation and class interaction than did late-stage students ($p < 0.01$). First-years especially found interaction with peers more helpful than mid-career ($p < 0.05$) and late-stage students ($p < 0.001$). In contrast, they rated individual work less helpful than did mid-career students ($p < 0.01$). Overall, first-year students in particular found active participation and peer interaction very helpful for learning mathematics.

Similarly, comparison of students' course experiences by the number of prior college mathematics courses helps to explain results from learning gains. Unlike non-IBL students, inexperienced IBL math-track students rated interaction as more helpful than did the most experienced students (6 or more courses, $p < 0.05$). Less experienced students also rated their own active participation more helpful than did the most experienced IBL students, although this difference was not statistically significant. Together these results suggest why younger and less experienced students report higher learning gains from IBL course practices (Section 3.2.4, 3.2.5) than do their older or more experienced classmates.

3.3.5 Other differences in course experiences were minor.

We also investigated whether course experiences depended on students' other demographic and academic characteristics, including GPA, ethnicity, and academic major.

In the analysis by GPA, there was no overall difference by GPA level in the course experiences reported by IBL or non-IBL math-track students. However, within the medium-achieving GPA group (GPA 3.0-3.79) separately, there were clear differences between the experiences of IBL and non-IBL students. The IBL students rated their own active participation, interactions, and assignments other than tests as more helpful than did their non-IBL peers ($p < 0.05$).

Similarly, the high-achieving math-track IBL students (GPA ≥3.8) rated the usefulness of other assignments higher than did non-IBL students. But they considered individual work less helpful than did their non-IBL peers ($p < 0.05$). Among low-achieving students, similar differences appeared but were not statistically significant. In sum, IBL students at varying achievement levels reported greater help from course practices than did their peers in traditional math classes. This finding was strongest among the largest group, medium-achieving IBL students. The most helpful practices were students' own active participation, interactions, and non-test assignments. However, these results do not clearly explain the observed differences in learning gains by students' GPA level (Section 3.2.6).

In the analysis by ethnicity, the number of Hispanic students was very small; no clear difference appeared in their course experiences versus those of non-Hispanic classmates. However, Hispanic IBL pre-service teachers found individual work slightly more helpful than did their non-Hispanic peers ($p < 0.05$). Consistent with our results on learning gains, Asian IBL students reported slightly more positive course experiences than white IBL students. This applied to tests, individual work, and other assignments, but these differences were not statistically significant.

Academic major was not a good predictor of IBL students' course experiences. However, IBL math majors' course experiences were clearly more positive than non-IBL math majors. They found the overall instructional approach more helpful than did their non-IBL peers ($p < 0.01$). They also gave higher ratings to their own active participation ($p < 0.01$), other assignments than tests ($p < 0.01$), and interaction with others ($p < 0.001$). Within the smaller major groups, science IBL majors found interactions more helpful than did non-IBL science majors. Consistent with their higher learning gains (Section 3.2.7) the small group of IBL economic majors had more positive class experiences than their non-IBL peers.

3.3.6 Summary of Findings on Students' Learning Gains and Class Experiences

The survey results on self-reported learning gains and experiences of instructional practices show that IBL courses were clearly beneficial for students' learning of mathematical knowledge, thinking and problem-solving. IBL classes also built up students' confidence, positive attitude, persistence and their ability to work with other students more than did the comparative non-IBL sections. IBL math-track students reported more varied and higher cognitive, affective and social gains than did non-IBL math-track students. IBL pre-service teachers reported lower gains than IBL math-track students. IBL students also described changes in their ways of learning mathematics and approaching problems more than did non-IBL students.

The difference between active IBL class practices and traditional lecture-based instruction was also apparent in students' experiences and learning gains. Unlike in lecture-based courses, class assignments other than tests, together with their own active participation and interactions with peers and instructors, contributed to IBL math-track students' learning gains. Their reports of "what helped me learn" emphasize everyday classroom experiences. Pre-service teachers' IBL class experiences differed from that of IBL math-track students. For non-IBL students, math tests played an important role in their class experiences and learning gains, suggesting that in those courses, student learning opportunities centered on tests and test preparation.

We also studied self-reported learning gains and class experiences within various subgroups. The most important finding related to math-track women, where IBL instruction may have an important role in diminishing gender differences in mathematics:

- IBL math-track women reported higher gains and more clearly positive class experiences as compared to non-IBL math-track women. This difference was independent of women's prior achievement in mathematics.
- Unlike in traditional lecture-based classes, women benefited from IBL instruction equally or slightly more than men. IBL math-track men also reported more and higher gains than non-IBL math-track men, but the differences were somewhat weaker.

We found additional interesting differences in class experiences and learning gains depending on students' mathematical experience and achievement level:

- Younger and less mathematically experienced IBL students had more positive class experiences. They also reported higher gains than older and experienced IBL students.
- Differences in learning gains between IBL and non-IBL math-track students were most clear among first-year students and the mid-career group.
- IBL students at all achievement levels reported greater help from course practices than did their peers who took traditional math classes. This finding was strongest among the largest group, medium-achieving IBL students.
- Lower-achieving students benefited more from IBL courses in comparison to traditional instruction. In contrast, high-achieving students benefited more than lower-achieving students from traditionally taught courses.
- These higher learning gains were strongest for low-achieving pre-service teachers. Medium-achieving pre-service teachers also reported higher gains than their high-achieving IBL peers. As Chapter 5 will show, these self-report findings are corroborated by test results.

We also investigated class experiences and learning gains against students' race/ethnicity and academic major but found no major differences among IBL students.

- Students' race, ethnicity and major (math, science, engineering/CS, economics, or non-science) did not affect IBL students' class experiences or learning gains.
- Among mathematics majors, IBL students' course experiences and gains clearly exceeded those of non-IBL peers. Modest variation in experiences and gains between IBL and non-IBL sections was found for students of other academic majors.

3.4 Survey Findings about Students' Beliefs, Motivation and Strategies

Based on preliminary evidence that students enjoy IBL approaches and find them interesting, we wished to examine students' mathematical beliefs, affect, goals and strategies—both at the beginning of a mathematics course, and how these changed during their course. Here we present results based on pre-/post-survey data from IBL and non-IBL math-track students, IBL and non-IBL pre-service teachers. The non-IBL pre-service teachers were mathematics majors pursuing a high school teaching certificate and thus are more comparable to other math-track students than to IBL pre-service teachers who were taking mathematics courses targeted to K-12 teaching. Details of the survey and items are presented in Appendix A3.

3.4.1 "Attitudinal" survey variables characterize students' beliefs, motivations, and mathematics learning strategies before and after their mathematics course.

A set of "attitudinal" variables is used to characterize students' mathematical beliefs, motivation and strategies for learning and problem-solving, by student group. Each attitudinal variable is constructed from one or more questions related to a common theme, previously known to be important from mathematics education research.

'Motivation' variables address students' reasons for learning mathematics as an academic subject and as a personal interest pursued on one's own time. The 'goals' variables address the balance of intrinsic, or internal, learning goals, against extrinsic goals such as getting good grades.

Novice views of the discipline are reflected in two variables: a view of problem-solving as a matter of repeated practice; a view of proofs as confirming knowledge rather than constructing it. These contrast with more expert-like views of problem-solving as applying rigorous reasoning, and proving as constructing knowledge.

'Beliefs on learning' variables contrast views of learning as instructor-driven with views that value collaborative learning and exchange of ideas. Self-regulation is an important skill in problem-solving, as it involves planning and assessing one's own thinking, and thus the ability to be flexible in approaching problems and to recognize when to abandon unproductive approaches. This variable addresses students' use of productive, self-regulatory strategies in learning. The use of independent and collaborative strategies is also important in solving mathematics problems.

Students' interest in teaching is probed by three variables: motivation to teach mathematics; a goal of communicating mathematics; and confidence in one's ability to teach mathematical ideas to others. Table 3.5 summarizes students' initial ratings on all these attitudinal variables.

3.4.2 IBL and non-IBL math-track students entered courses with similar beliefs, motivations, and mathematics learning strategies.

As Table 3.6 shows, IBL and non-IBL math-track students held similar attitudes at the beginning of the courses we studied. Among math-track students, both IBL and non-IBL students (including classmates of the latter group who were pre-service teachers) started with high motivation for studying mathematics. They had strong intrinsic and extrinsic goals and high confidence in their ability to do and teach mathematics. Generally they preferred instructor-driven instruction but also valued exchanging ideas with other students. Both practice and rigorous reasoning were important for their problem-solving. These advanced students had reasonably sophisticated ideas about proving: most tended to consider proving to be a constructive activity, a tool for understanding mathematical ideas rather than something to be practiced or recalled.

Table 3.6: Average Beliefs, Motivation and Strategies at the Start of a Course (pre-survey)

Attitudinal variable		Average rating (7-point scale)				Sig. level (for each pairwise comparison)
		math-track students		**pre-service teachers**		
		A IBL (N=570)	**B** Non-IBL (N=375)	**C** IBL (N=233)	**D** Non-IBL (N=25)	
Motivation	*personal interest*	4.6	4.2	3.3	4.6	C < A, B***, D**; B < A**
	math major	4.7	4.3	3.5	5.7	C < A***, B, D**; D < B*
	math future	6.0	6.2	5.5	6.1	C < A, B***
	teaching	3.9	3.7	6.1	6.0	A, B < C, D ***
Goals	*intrinsic*	5.8	5.8	5.3	5.6	C < A, B***
	extrinsic	5.4	5.5	5.9	5.7	A***, B** < C
	communicating	5.3	5.2	6.1	6.1	A, B < C***, D*
Enjoyment		5.4	5.2	4.2	5.4	C < A, B, C***
Confidence	*math ability*	5.1	4.8	5.0	5.2	B < A*
	teaching	5.4	5.3	5.2	5.8	
Beliefs about learning	*instructor-driven*	5.3	5.3	5.6	5.4	A**, B* < C
	group work	4.6	4.2	4.8	4.9	B < A**, C***
	exchange of ideas	5.2	5.0	5.3	5.3	B < C*
Beliefs about problem-solving	*practice*	5.0	5.2	5.3	5.3	
	reasoning	5.3	5.2	5.2	4.9	
Beliefs about proofs	*constructive*	5.7	5.6	5.5	5.6	
	confirming	4.7	4.8	5.1	4.9	A < C**
Strategies	*independent*	5.3	5.3	4.9	5.1	C < A, B***
	collaborative	4.6	4.5	5.0	5.0	A**, B*** < C
	self-regulatory	5.2	5.2	5.3	5.0	

p < .05, p < .01**, p < .001****

Compared with pre-service teachers, who more commonly emphasized collaboration, math-track students emphasized individual learning and problem-solving strategies. All the students reported a high level of self-regulatory activities, referring to their ability to plan and monitor their mathematical work.

The differences between IBL and non-IBL math-track students were minor. However, IBL students reported slightly higher interest in mathematics as a personal, not just academic interest: reading magazines or articles related to math, participation in a math-related club, and bringing up mathematical ideas in a non-mathematical conversation ($p < 0.01$). Also, IBL students preferred group work more than non-IBL students ($p < 0.01$) and reported higher math confidence ($p < 0.05$). These findings show that there is some self-selection of students into active and collaborative IBL courses. Students who enrolled in IBL courses were personally more interested in mathematics, displayed slightly higher confidence in their mathematical ability, and had a more positive view about group work as compared to other math-track students. This is an important finding that cannot be ignored in interpreting other results.

3.4.3 Pre-service teachers started with high interest in teaching but less positive attitudes.

Pre-service teachers were a distinct group whose motivations, learning beliefs, and strategies for learning and problem-solving differed from those of math-track students. These students found less enjoyment ($p < 0.001$) and were less interested ($p < 0.001$) in studying mathematics. Such views are typical of elementary or middle school pre-service teachers (Harpe & Daane, 1998; MacNab & Payne, 2003). However, the commonly reported lack of confidence in their own mathematics abilities (Austin, Wadlington, & Bitner, 2001; Brady & Bowd, 2005) did not show up strongly among this group of pre-service teachers—possibly because these campuses generally attract well-prepared students.

Compared with math-track students, IBL pre-service teachers reported stronger interest in teaching ($p < 0.001$) and stronger goals toward communicating ($p < 0.001$). They also preferred traditional instruction, as seen their orientation toward instructor-driven practices ($p < 0.05$) and extrinsic goals for learning ($p < 0.01$). They reported more traditional beliefs about the nature of mathematical proofs as compared to IBL math-track students ($p < 0.01$).

The number of non-IBL pre-service teachers was low (25), so comparing them with other groups is less reliably meaningful. Their higher interest ($p < 0.01$) and enjoyment of studying mathematics ($p<0.001$) in comparison to IBL pre-service teachers are explained by the fact that they all were math majors taking the same advanced courses as their non-teaching peers. Their beliefs, motivation and strategies were like those of general math-track students. But, like IBL pre-service teachers, they rated goals for teaching as more personally important than did math-track students ($p < 0.05$).

Overall, the pre-survey results show that the students in this study were academically well motivated. Observed patterns of difference among the groups are consistent with prior studies and with instructors' comments on their students.

3.4.4 Only minor changes were observed in students' beliefs, motivation, and strategies.

After probing student beliefs, motivation, and problem-solving strategies in the pre-survey, we compared these to post-survey data from 754 individually matched surveys. Despite the high initial averages for many variables, some pre/post changes could nonetheless be seen. However, these were often minor changes that were not statistically significant.

We studied the changes within the three student groups: IBL math-track students, non-IBL math-track students, and IBL pre-service teachers. As Table 3.7 shows, each of the three student groups responded slightly differently to their mathematics course.

Table 3.7: Average Changes in Beliefs, Motivation and Strategies (Pre- vs. Post-Survey)

Attitudinal variable		IBL math-track (N=390)	Effect size	Non-IBL math-track (N=156)	Effect size	IBL pre-service teachers (N=184)	Effect size
Motivation	*personal interest*	+	0.11	+			
	math major						
	math future	-	0.12	- -	0.24	- - -	0.41
	teaching					+	
Goals	*intrinsic*						
	extrinsic	+	0.12	+		-	0.15
	communicating	+ +	0.25	+			
Enjoyment				-			
Confidence	*math ability*			- -	0.33		
	teaching	+		+			
Beliefs about learning	*instructor-driven*			+		-	0.18
	group work	+	0.25	+		+	0.18
	exchange of ideas	+	0.15	+		+	
Beliefs about problem-solving	*practice*					-	
	reasoning			-			
Beliefs about proofs	*constructive*	+	0.14			+	0.22
	confirming					+ +	0.19
Strategies	*independent*					+	
	collaborative	+ +	0.24	-		+	
	self-regulatory			-		- - -	0.36

slight change (+, -), stronger change (+ +, - -), clear change (+ + +, - - -)

Here we report the impact of taking the course in the form of effect size, a measure of the impact of the course based on the magnitude and direction of the change of the post-survey response as compared to the pre-survey. However, the effect sizes for the statistically significant changes varied only between 0.11 and 0.41, which are generally small effects. That is, both IBL and non-IBL courses had only minor effects on students' beliefs, motivation and strategies.

3.4.5 IBL and non-IBL courses had opposite effects on math-track students' confidence and collaboration.

Among math-track students, most of the changes recorded by IBL students were slight increases, whereas non-IBL students' ratings more often reflected slight decreases. Non-IBL students reported several changes that suggest negative effects of their courses:

- weakened confidence in their own math ability
- decreased willingness to study hard for and take additional math courses
- reduced enjoyment of mathematics
- less strong belief in rigorous reasoning as a general problem-solving approach
- lower use of collaborative and self-regulatory strategies in learning and problem-solving.

However, non-IBL students also reported some attitude-related increases, some of which appear contradict one another:

- stronger beliefs in instructor-driven teaching, but also in group work and exchange of ideas with others
- stronger extrinsic learning goals, but also stronger goals for communicating mathematics
- slightly stronger confidence in teaching mathematics.

Among IBL students, the observed changes suggest that their course had largely positive effects:

- stronger personal interest in mathematics, extrinsic goals, and goals for communicating about mathematics
- stronger beliefs in group work and exchange of ideas with other students
- stronger belief in mathematical proving as a constructive activity
- increased use of collaborative learning and problem-solving strategies
- slightly greater confidence in teaching mathematics.

However, their willingness to study hard for and take additional math courses slightly decreased.

The differences between IBL and non-IBL math-track students' changes in two areas—mathematics confidence ($p< 0.05$) and use of collaborative strategies ($p< 0.01$)—were statistically significant. These findings suggest that, unlike IBL students, students in a traditional math course tended to lose some confidence in their ability to do mathematics. After lecture-based instruction, non-IBL students also less frequently collaborated with other students while solving problems. In addition, they believed slightly less in the importance of rigorous reasoning in studying college mathematics.

In contrast, after an IBL course, math-track students were more prone to use collaborative learning and problem-solving strategies. They believed more in the benefit of working with other students. Moreover, their priority on communicating about mathematics increased more than among non-IBL students. Although both non-IBL and IBL math-track students' willingness to take additional math courses slightly decreased, this was more true for non-IBL students.

Overall, the results suggest that IBL math courses tended to promote more sophisticated and expert-like views of mathematics and more interactive approaches to learning. In contrast, traditional mathematics courses appeared to weaken students' confidence and enjoyment, and did not help them to develop expert-like views or skillful practices for studying college mathematics.

3.4.6 IBL instruction had mixed effects on pre-service teachers' beliefs and behavior.

IBL pre-service teachers started with weaker motivation to study mathematics, although they had higher interest in teaching. The effects of IBL classes on their beliefs, motivation, and strategies were somewhat mixed. After an IBL course, pre-service teachers:

- were less willing to take additional math courses or study hard for mathematics (d=0.41).
- emphasized less extrinsic learning goals (d=0.15)
- preserved but did not gain confidence in their own math ability
- preferred group work more strongly (d=0.18) and instructor-driven instruction less strongly (d=0.18)
- strengthened both constructive (d=0.22) and traditional (d=0.19) views of mathematical proofs
- less often used self-regulatory strategies in learning (d=0.36).

Weakened emphasis on extrinsic goals and instructor-driven activities suggest that students are developing more mature approaches to learning mathematics. IBL pre-service teachers' (already strong) goals for communicating about mathematics and use of collaboration increased less than among IBL math-track students. However, the decline in their use of important self-regulatory strategies for learning mathematics was different from the trend for IBL math-track students ($p < 0.001$). The decline in their willingness to take additional math courses is explained by the fact that they did not need to study more mathematics. That is, completion of the targeted, two-course IBL sequence satisfied their mathematics requirement, and additional mathematics courses were unlikely to fit within their already-full academic program.

3.4.7 Math-track women gained confidence and motivation during IBL classes.

Generally, the attitudinal changes both for IBL math-track women and men were minor. Both groups largely preserved their high motivation, supportive mathematical beliefs, and strategies. However, only women slightly grew in confidence in their own math ability and in teaching mathematics. Women also grew more than men ($p < 0.01$) in their preference for group work.

Among math-track women, we found an interesting difference between IBL and non-IBL women's development of confidence ($p < 0.01$). For non-IBL women, confidence in their own mathematics ability clearly dropped during a traditional course ($p < 0.01$, $d=0.51$). In contrast, IBL women's math confidence tended to increase during their course. Moreover, their confidence in teaching increased slightly more than did that of non-IBL women. We also observed stronger decreases in non-IBL women's interest and motivation to study mathematics, as compared to IBL women. These results strengthen our findings that traditional ways of teaching tended to weakn women's confidence and motivation to study mathematics, whereas an IBL instructional approach strengthened their confidence and motivation.

3.4.8 Traditional instruction had negative effects on women's beliefs and learning strategies.

We also saw changes in non-IBL math-track women's beliefs and practices that are unsupportive of their development as mathematicians: declines in beliefs about math learning (e.g., less group work or exchange of ideas with other students), in their beliefs about math problem-solving (e.g., less rigorous reasoning or multiple approaches), and in their use of collaboration to solve problems ($d=0.42$). These contrast with slight increases in the same areas for IBL math-track women. In addition, IBL women more than non-IBL women increased their preference for group work ($p < 0.01$), an efficient problem-solving strategy.

3.4.9 First-year students developed learning-enhancing beliefs and strategies during IBL classes.

There were also some differences in the changes seen for less and more experienced students. Among younger students, IBL experiences helped to develop efficient beliefs about and strategies for mathematics learning and problem-solving, though they did not enhance these students' personal interest in mathematics. First-year students' beliefs about mathematical problem-solving shifted towards a view of problem-solving as a constructive ($d=0.27$) rather than confirming activity. They made more use of independent thinking and problem-solving ($d=0.20$), and slightly more use of collaborative strategies. However, their teaching confidence and personal interest in mathematics slightly decreased ($d=0.20$).

3.4.10 Sophomores and juniors gained interest and motivation during IBL classes.

For mid-career students (sophomores and juniors), an IBL course seemed to have a positive impact on personal interest in mathematics ($p < 0.05$, $d=0.18$) and motivation to study mathematics in the future ($p<0.01$, $d=0.23$). They also preserved or increased their high goals for learning math, and also their confidence. The gain in interest, particularly, contrasts with the observations for the two other student groups.

However, mid-career students displayed slightly more traditional views about math learning and problem-solving at the end of an IBL course. Unlike first-years, they were more convinced at the end of an IBL course that proofs were truths to be confirmed rather than constructed on the basis of their own understanding ($p < 0.05$). They also emphasized slightly less than first-year students

($p < 0.05$) independence in learning and problem-solving. These results suggest that an IBL experience did not help in changing mid-career students' beliefs about mathematics, but did modestly increase their interest in and motivation to study mathematics.

3.4.11 IBL classes strengthened late-stage students' confidence.

As might be expected, no substantial changes were seen in beliefs, motivation or strategies of late-stage students (senior and older). It is harder to change advanced students' beliefs and study habits from those that have been successful for them so far. However, a few interesting things still came out. Unlike with the two other groups, these students' motivation to teach mathematics declined ($p < 0.01$, $d=0.28$), but their confidence in teaching mathematics and general math confidence slightly increased. At the end of an IBL course, they also more strongly believed in multiple approaches and rigorous reasoning in solving challenging math problems ($p < 0.05$, $d=0.27$). Their use of self-regulatory strategies increased slightly.

3.4.12 Traditional teaching yielded some negative effects on low, middle and high achievers.

We also studied the changes in beliefs, motivation and strategies for different achievement groups, established on the basis of self-reported GPA. These results mostly concern sophomores and older students, as first-year students could not report a current GPA. Generally, no clear differences were found between low, middle- and high-achieving IBL math-track students. Comparison of the changes for IBL and non-IBL students show that, at the end of their course:

- Low-achieving non-IBL students' overall math confidence decreased (although their teaching confidence increased).
- Unlike IBL students, non-IBL middle-achievers lost some personal interest in studying mathematics. Moreover, their beliefs about proofs turned towards less expert-like views.
- Middle-achieving non-IBL students used less effective problem-solving strategies after a traditional course, whereas middle-achieving IBL students reported increased use of effective strategies, in particular self-regulatory strategies.
- High-achieving non-IBL students' confidence slightly dropped during a traditional course, whereas confidence increased among equivalent IBL students.
- High achievers preferred instructor-driven teaching more clearly at the end of a lecture-based math course than after an IBL course.

Together, these findings indicate that lecture-based instruction does not encourage—and can harm—students' confidence, use of effective problem-solving approaches, and development of expert-like views of mathematics. The effects may depend somewhat on students' level of prior college success.

3.4.13 Learning gains were modestly connected to changes in students' beliefs, motivation, and strategies.

We looked to see whether the observed changes in students' beliefs, motivation and strategies were connected to their learning gains. Table 3.8 displays the statistically significant correlations found for IBL math-track students. Changes in students' beliefs, motivation and strategies were not large and thus these changes did not explain much of the variation in IBL math-track students' learning gains.

Table 3.8: Statistically Significant Correlations between Changes in Beliefs, Motivation and Strategies and Learning Gains, for IBL Math-track Students

Attitudinal Variable		Correlation with Learning Gains			
		Math concepts & thinking	Application	Affective	Social
Motivation	*interest*	0.159**	0.108*	0.158**	0.088
	math major	0.204**	0.142**	0.188**	0.120*
	math future	0.117*	0.113*		
	teaching	0.154*		0.110*	
Goals	*intrinsic*	0.111*		0.136**	
	communicating	0.116*		0.105*	
Enjoyment		0.213**	0.163**	0.212**	0.136**
Confidence	*math ability*	0.224**		0.185*	
Beliefs about learning	*group work*	0.193**	0.178**	0.192**	0.223**
	exchange of ideas	0.202**	0.193**	0.160**	0.109*
Beliefs about problem-solving	*reasoning*	0.162*	0.153*		
Beliefs about proofs	*constructive*	0.125*			
Strategies	*independence*	0.176**		0.104*	0.155**
	collaborative	0.117*	0.161**	0.114*	0.263**
	self-regulation	0.142**		0.113*	0.143**

* $p < 0.05$, ** $p < 0.01$

The correlations in Table 3.8 do show that IBL math-track students who reported higher cognitive, affective, and social learning gains also reported increased motivation, enjoyment and math confidence. We suggest that, as students had successful learning experiences with IBL methods, they also developed stronger beliefs in the value of group work and exchange of ideas with other students. Students' increased use of independent, collaborative and self-regulatory strategies in learning also contributed to their learning gains. Furthermore, cognitive gains were reinforced by strengthened beliefs in rigorous reasoning and constructive activity in proving.

While the linkages are fairly modest in strength and fairly broad across the attitudinal variables, they are not entirely generic. Results from regression analysis show that students' gains in learning mathematical concepts and thinking were best explained (8%) by an enhanced preference for exchanging ideas with others (β= 0.216) and increased interest in graduating with a math major (β= 0.159). (Here β is a coefficient that indicates the relative importance of each factor in explaining the learning outcome.) Students' affective gains were best explained (7%) by increased enjoyment of mathematics (β= 0.141), strengthened value on group work (β= 0.134), and increased interest in graduating with a math major (β= 0.129). Gains in collaboration were explained by (8%) increased use of collaborative strategies (β= 0.202) and strengthened value of group work (β= 0.144). Increases in mathematical confidence explained only 3-4% of IBL math-track students' gains but 13-16% of IBL pre-service teachers' learning gains.

For IBL pre-service teachers, gains in mathematical concepts and thinking were best explained (9%) by increased value on rigorous reasoning (β= 0.211) and increased enjoyment of mathematics (β= 0.187). Increased enjoyment in particular explained their affective gains (14%; β= 0.344), together with increased belief in exchange of ideas as valuable for learning mathematics (β= 0.157). Their gains in collaboration were linked (11%) to increased use of self-regulatory strategies (β= 0.262) and strengthened value on group work (β= 0.169). Overall, the correlations point to somewhat specific relationships between particular attitudinal changes and particular types of learning gains.

3.4.14 Summary of Findings on the Nature of Students' Beliefs, Motivation and Learning Strategies and Changes in These after College Mathematics Classes.

Overall, the results presented in Section 3.4 show that both IBL and non-IBL students had rather high motivation and confidence at the beginning of their mathematics courses. They valued traditional teaching slightly more than group work, but applied both independent and collaborative strategies in their own learning and problem-solving. They held constructive views about proving and valued both practice and rigorous reasoning as tools in solving problems. Compared to math-track students, IBL pre-service teachers started with higher interest in teaching but more negative attitude toward mathematics.

Changes in students' beliefs, motivation, and mathematics learning strategies as a result of their course were modest for all students. However, we observed some negative effects of traditional teaching on non-IBL math-track students' confidence and collaboration. In contrast, IBL math-track students valued and used more collaboration in learning after their course. For women, traditional teaching weakened confidence and motivation while IBL instruction strengthened these. Moreover, traditional instruction had negative effects on women's beliefs about learning and mathematics and on their use of effective learning strategies.

There were some modest positive shifts in the attitudes of pre-service teachers who took an IBL course. They believed more in group work and less in lecture-based teaching, although they were less likely to use self-regulation in learning mathematics after their IBL course. They placed less emphasis on extrinsic goals such as grades and earning a degree. We saw no change in the math confidence of pre-service teachers during an IBL course. But given the confidence declines observed in women taking math-track courses (and since most of the pre-service teachers are women), this may be a more positive result than it appears on the face. Without comparative sections of non-IBL pre-service courses, it is difficult to interpret this finding.

Overall, although the attitude-related survey data provide less strong findings than does the survey of learning gains, they show some important positive effects of IBL instruction on students' views of mathematics and approaches to learning. They also reveal some significant negative consequences of traditional teaching, especially on women's mathematics learning.

IBL approaches helped first-year students to develop enhancing beliefs and strategies for learning and problem-solving, sophomores and juniors to gain in interest and motivation, and senior and more advanced students to strengthen their confidence. These results indicate varied but positive effects of IBL instruction on students' learning at all stages in the college career. Changes in IBL math-track students' beliefs, motivation and strategies only partially explained their learning gains. A growing preference for group work and exchange of ideas with others—activities that were characteristic of IBL classes—contributed to students' cognitive, affective, and social learning gains. In particular, these and the use of collaborative strategies enhanced students' gains in collaboration. Learning gains were also supported by increased interest, enjoyment of mathematics, and enhanced confidence in math ability.

3.5 Conclusion, Strengths and Limitations of the Survey Studies

3.5.1 Strengths of the Survey Studies

Several features ensure that our results from the attitudinal and learning gains survey can be trusted. Relative to most previous studies, the numbers of courses and students from four campuses are large, making our survey findings reliable and applicable somewhat more generally. Moreover, the large variety of IBL instructional practices used in the sections enriches our survey data and strengthens the findings. The results generally favoring IBL courses are consistent despite the variety of active classroom practices and student audiences within the IBL sections. This diversity of courses, students, and teaching approaches is reflective of real-world educational reform, rather than an idealized laboratory study—thus the conclusions should apply to other real-world efforts to improve teaching and learning in college mathematics departments.

Our findings are based on careful study and survey design and data-gathering. The attitudinal survey questions reflect important aspects of learning known from previous research in mathematics education: motivation, enjoyment, confidence and beliefs about mathematics learning and problem-solving. Items and sections of the survey were tested and revised prior to wider use. Our learning gains post-survey (SALG-M) was based on a well-validated instrument that confirmed validity and reliability of the findings. The survey questions are concrete and student responses are anonymous, factors that help to produce accurate self-assessments from students (Seymour, Wiese, Hunter, & Daffinrud, 2000). Using self-report enables us to include a large and diverse sample, because (unlike tests) the gains measures are not course-specific

The quality of our survey data was also high. Firstly, the large number of completed surveys ensured the reliability of our findings. Separate analyses using standard statistical measures of reliability showed that the instruments produced real and non-arbitrary results, with good reliability scores on nearly all survey items and acceptable scores on the six least reliable composite variables.

We also were careful and thorough in gathering and dealing with the survey data. All the data were gathered anonymously, which ensured that students could give candid responses, and using standardized procedures to provide enough time to complete the questions. Students could also choose not to participate in the study.

Correlational analysis showed that the composite variables had a good construct validity. They produced real results on students' learning gains, motivation, beliefs and strategies. In addition, consistent findings from multiple data sources (open-ended survey questions, student interview data) strengthen the validity and reliability of the survey results.

3.5.2 Limitations of the Survey Studies

The limitations of the survey study are related primarily to limitations in the samples. Our samples represented academically rather good students, especially among IBL math-track students. All students reported high motivation and rather sophisticated mathematical beliefs and strategies. Changes in student attitudes or learning and problem-solving strategies are difficult and usually take time. Among good students, such changes may be even less apparent. Students also reported quite high learning gains. While we observed differences in learning gains and development of attitudes between IBL and non-IBL students, there are indications in the data that such differences may be even stronger among weaker and less experienced students. This was reflected, for example, in the clearer benefit of IBL instruction to lower-achieving students as compared to others. Thus, our samples may not be optimized to detect the results of IBL instruction.

Because many instructors of non-IBL sections chose not to have their students participate, the sample of math-track non-IBL students was smaller than ideal. This made comparisons of attitudes and learning gains between various student subgroups more difficult and produced less reliable results on group differences. The non-IBL sample also largely consisted of first-year students taking calculus, while the IBL sample included high numbers of older and later-stage math-track students. Our data on pre-service teachers' gains and attitudes came mostly from women who were taking targeted IBL courses for future elementary/middle school teachers, with much less data from secondary school IBL pre-service teachers and little data from men. No courses were available for making real comparisons between IBL and non-IBL pre-service teachers.

Finally, we were unable to match all the pre- and post-survey data on individual level due to missing or mismatched identifiers in some surveys. These items were most often absent from surveys completed by non-IBL students who participated in large, lecture-based courses. Some of the students also found the combined post-survey too long and did not fill in all the questions.

We had few opportunities to get direct information about students' performance in these courses, such as course or exam grades or samples of student work. We were thus largely unable to study the direct relationship between students' attitudes or self-reported learning gains and the development of their mathematical knowledge and skills as measured by tests. We did have some test data sets (Ch. 5) and instructor ratings of math-track students at two campuses, but both types of data sets were small and we were mostly unable to match these results with individual students' responses to the surveys.

3.5.3 Conclusions

IBL teaching methods that emphasize students' own active participation and collaboration with other students create a context different from traditional college mathematics instruction. Active engagement of students in their own learning processes, with responsibility, collaboration, and efficient use of personal resources, is seen to enhance growth of students' thinking and problem-solving (Prince & Felder, 2007) and social skills (Duch et al., 2001; Jordan & Metais, 1997). Overall, our results indicate that use of active instructional methods with inquiry and collaboration provided a learning context that had powerful effects on students' learning of college mathematics. Students' own active participation and interactions in class particularly helped their learning in IBL classes. In IBL sections, non-test assignments and homework were clearly beneficial for student learning.

Overall, IBL math-track students reported more positive experiences and higher cognitive, affective and social gains than did non-IBL math-track students at the end of a course. Math-track students began with high motivation and adequate beliefs about mathematics learning and problem solving. However, participation in an IBL mathematics course further promoted these views and approaches. IBL courses especially changed students' beliefs and practices towards enhanced collaboration. They helped students to preserve or gain confidence to do mathematics and somewhat increased their enjoyment of and personal interest in mathematics. These results contrasted with the outcomes from, lecture-based courses, which appeared to have some negative effects on students' beliefs and problem-solving strategies. Our results thus confirm that various forms of IBL instruction may be especially efficient in promoting college mathematics students' beliefs (cf., Kwon, Rasmussen, & Allen, 2005; Smith, 2006).

IBL pre-service teachers' lower interest in studying mathematics but higher emphasis on collaboration and teaching were reflected in their attitudes and beliefs. They too responded positively to IBL class activities but reported lower gains and less benefit from IBL learning approaches than did math-track students. However, their high gains in applying mathematical knowledge, collaboration, and comfort in teaching mathematics have an important effect on their future teaching of mathematics. Moreover, they more than other IBL students reported changes in their ways of learning mathematics and solving problems. They developed more mature beliefs and approaches to learning mathematics during an IBL course. Previous research suggests that pre-service teachers' own beliefs and learning experiences importantly influence the ways that they teach mathematics in their own classes (Ma, 1999; Raymond, 1997; Thompson, 1992). The positive outcomes from IBL classes may thus have significant contributions to how these future teachers will later perform in K-12 classrooms.

IBL instruction appeared most beneficial for math-track women, and thus suggest that IBL approaches may have an important role in diminishing gender differences in mathematics. While both men and women in IBL sections reported stronger learning gains than their non-IBL peers, IBL women reported as high or higher gains than IBL men. In contrast, women in non-IBL sections reported lower cognitive and affective gains than men. Moreover, traditional ways of teaching seemed to weaken women's confidence and motivation to study mathematics, whereas an IBL approach strengthened women's confidence and motivation. Also math-track men clearly benefited from IBL instruction as compared to traditional lecture-based teaching. This appeared in IBL math-track men's more positive experiences and higher gains than that of non-IBL math-track men. However, unlike among women, these differences were also related to IBL men's higher prior or expected achievement level as compared to that of non-IBL men.

We found some interesting differences in attitudes, class experiences and learning gains between various student sub-groups. Younger and less experienced math-track IBL students had more positive class experiences and reported higher gains than older and experienced IBL students. The differences in learning gains between IBL and non-IBL math-track students also were most clear among first-year students and the mid-career group. Moreover, lower-achieving students benefited more from IBL classes in comparison to traditional instruction. In contrast, high-achieving students benefited more than lower-achieving students from traditional teaching of college mathematics. While late-stage and highest achieving mathematics students benefit from IBL instruction, the benefits may be strongest among lower achieving and younger or less experienced students. Some of the results in Chapters 5 and 6 bolster these findings.

3.6 References Cited

Austin, S., Wadlington, E., & Bitner, J. (2001). Effect of beliefs about mathematics on math anxiety and math self-concept in elementary teachers. *Education, 112*(3), 390-396.

Brady, P., & Bowd, A. (2005). Mathematics anxiety, prior experience and confidence to teach mathematics among pre-service education students. *Teachers and Teaching* 11(1), 37-46.

Capraro, R. M, Capraro, M. M., Parker, D., Kulm, G., & Raulerson, T. (2005). The mathematics content knowledge role in developing preservice teachers' pedagogical content knowledge. *Journal of Research in Childhood Education, 20*(2), 102-118.

Harper*, &* N. J., & Daane, C. J. (1998). Causes and reduction of math anxiety in preservice elementary teachers. *Action in Teacher Education, 19*(4), 29-38.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371-406.

Ma, L. (1999). *Knowing and teaching elementary mathematics.* Mahwah, NJ: Lawrence Erlbaum Associates.

MacNab, D. S., & Payne, F. (2003). Beliefs, attitudes and practices in mathematics teaching: Perceptions of Scottish primary school student teachers. *Journal of Education for Teaching, 29*(1), 55-68.

Raymond, A. (1997). Inconsistencies between beginning elementary teachers' beliefs and practice. *Journal for Research in Mathematics Education,* 28, 550-576.

Seymour, E., Wiese, D., Hunter, A. & Daffinrud, S. M. (2000). *Creating a better mousetrap: On-line student assessment of their learning gains.* Paper presented at the National Meeting of the American Chemical Society, San Francisco, CA, April 2000.

Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 127-146). New York: Macmillan.

Wilkins, J. L. M. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs, and practices. *Journal of Mathematics Teacher Education, 11*(2), 139-164.

Chapter 4: Findings on the Relationship of Student Gains to Classroom Activities

4.1 Introduction

In Chapter 2 we compared IBL and non-IBL courses with respect to what happened during class time. The two types of classrooms were rather different in classroom activities, leadership roles, and other measures of student-centered approaches to teaching and learning. In Chapter 3 we extensively discussed the differences in cognitive and affective gains reported by IBL and non-IBL students. IBL students reported higher gains on both cognitive and affective measures than their non-IBL peers. While these two findings suggest a connection between the way class is conducted and the gains students acquire from the course, this is circumstantial evidence only. In this Chapter, we explore direct connections between the classroom observation data and the student survey outcomes, addressing the research question:

- How do student learning outcomes relate to the nature of the instruction they experience?

So far, we have used the labels "IBL" and "non-IBL" to gloss the broadly different approaches used in particular classes, comparing student results by these labels. This is like asking whether or not a new drug helps treat a medical condition. However, these labels hide a good bit of variation (see Section 2.3). Here, our approach is to link data from direct classroom observation to student gains reported on the SALG-M post-survey. In the drug analogy, this is like asking what dose of the drug is needed for any effect on the disease, and what dose has the best effects.

However, these two types of data pertain to two different units of analysis: the classroom observation data describe course sections, and the survey data describe individual students. To address this mismatch, we computed student gain averages for each course section. While that allows us to relate the classroom observation data to the student gain outcomes, it also obscures the large variability of student gains within each course section.

Moreover, the match between observation data sets and survey data sets was not perfect. We analyzed observation data for 43 course sections at three campuses, but unfortunately did not obtain post-surveys from seven of those sections. Likewise, we could not gather observation data from all sections that returned student post-surveys. The final data set for the combined observation/survey analysis includes 30 IBL and six non-IBL sections, with averages representing 670 students.

As we worked on this research question using standard statistical approaches, we found that the underlying relationships between instruction and student outcomes were detectable but clearly complex and multi-dimensional. We asked our collaborator Tim Weston to apply a modern statistical approach called Hierarchical Linear Modeling, which can account for multiple and nested levels of variability—in this case, students nested within courses, both of which vary. We present some initial results from the HLM approach that indicate a few of the important course-level predictors of student success. But we also caution that these findings are less well developed than those presented in other chapters, and more work remains to be done before our conclusions are firm. We view these findings as promising but not complete.

4.2 Construction of Observation Variables to Indicate Student-Centered Approaches

4.2.1 Separate classroom activity and leadership variables did not relate well to student gains.

We discussed the observation data we collected and variables we constructed in detail in Chapter 2 and Appendix A2. For this chapter, observation variables of interest are percentages of class time spent on various activities (i.e. lecture, discussion, group work) (which we refer to as “class activity” variables) and percentages of class time spent with certain individual or groups in the leadership role (faculty, student, whole class) (“class leadership” variables). As both of these measure the percentage of class time spent in different ways, all are “class time” variables.

The correlation analysis between the student gains variables and individual class activity variables revealed small correlations that were not statistically significant. For example, we did not find a connection between the percentage of class time spent on discussion and student gains in mathematical thinking. We found no relationship between other pairs of student gains and class activity variables.

This result is not unexpected. Each instructor constructs his or her own blend of instructional activities, with differing emphases. In non-IBL sections, lecture was the most prevalent classroom activity and instructors played the leadership role most of the time. While generally IBL sections focused more on student-centered activities like discussion and group work, the specific ratio of different activities varied widely from classroom to classroom (see Section 2.3). Moreover, in any classroom, all these activities competed for class time: if more time was spent on lecture, there was less time left for discussion. Thus, there is high variability in the percentage of time spent on any activity, even among courses that emphasize student-centered methods. Hence, it is not surprising that individual activities did not correlate well with the student gains.

4.2.2 The sum variables were suitable measures of overall “student-centeredness” of each course.

Combining the percentages of time spent on all student-centered instructional activities provides a robust composite measure of the overall student-centered time. This *total percentage of student-centered time*, as an indicator of focus on active student learning, should relate to the learning gains students acquire from taking the course. If active student participation facilitates student learning, as the literature suggests (see e.g. Prince, 2004, and Section 1.4), the student learning gains should also be related to the *total percentage of student-led time.* The distinction between student-centered and student-led time is discussed extensively in Section 2.2.3.

The composite variables reflect real variation in how students spent class time. As such, they reflect differing emphases on student-centered, active-learning approaches—including variations such as interactive lecture or some IBL “experiments” that mixed lecture with student presentations. However, they do not capture the effectiveness with which instructors implemented these approaches. To address some aspects of implementation quality, we discuss observer survey ratings in Section 4.4.

4.3 Relationships between Student Gains and Classroom Time Variables

To look for statistical linkages between gains and classroom activities, we performed bivariate correlation analysis between the student gains variables and the four composite observation variables: total percentage of instructor-centered or student-centered activities, total percentage of instructor-led or student-led class time.

4.3.1 The relationship between total student-centered time and student gains was complicated and non-linear.

The total percentage of student-centered time correlated rather well with the student gain variables: most pairwise correlations were fairly large and statistically significant. Similarly, the total percentage of student-led time was rather highly related to the student gains outcomes, with large correlations that were mostly statistically significant. On the other hand, the percentages of instructor-centered and instructor-led time displayed highly negative and mostly statistically significant correlations with the student gains variables. The relationships between composite total percentage of time variables and selected student gain variables are presented in Table 4.1.

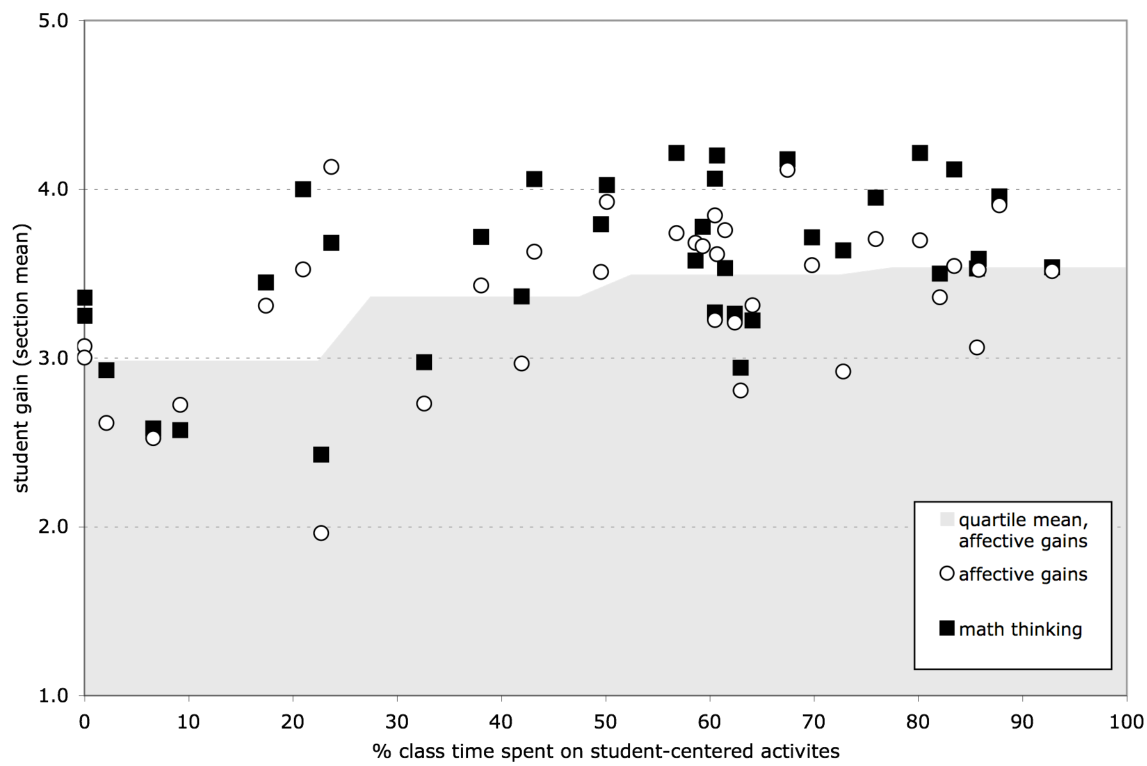
Table 4.1: Correlations of Section Means for Select Student Gains and Total Percentage of Time Variables

Student gains from survey	Total % of student-centered time	Total % of instructor-centered time	Total % of student-led time	Total % of instructor-led time
Mathematical thinking	0.385*	-0.402*	0.444**	-0.447**
Affective gains	0.371*	-0.383*	0.405*	-0.405*

Select correlations presented in Table 4.1 elucidate the general relationship very clearly (which was also seen for other gains not shown in the table). Student gains in mathematical learning and affective gains increased with increased time spent on student-centered activities. The direct relationship was even stronger for the rise of student gains with the total percentage of student-led time. On the other hand, the negative correlations for student gains with total instructor-centered and instructor-led time indicate an inverse relationship—as is expected, since student-centered and instructor-centered variables are near inverses. Overall, allotting more time for student-centered instructional activities and student leadership facilitated student learning, while spending more time on instructor-centered and instructor-led methods had a negative effect on student learning gains.

Although most of the correlations discussed above were rather large and statistically significant, some pairs of variables produced smaller correlations that were still statistically significant, and other were not significant. These results, and visual inspection of the scatter plots, suggest that the relationship is somewhat non-linear. But the correlation statistical test is specifically designed to evaluate the strength of a hypothesized linear relationship between two variables. For most variable pairs, some portions of the data followed a linear trend, which likely yielded the significant correlation test results. Other portions of the data, however, did not conform well to the linear approximation. Overall, the linear correlation test returned weaker outcomes for the more non-linear relationships. Figure 4.1 shows the scatter plot for the two survey gains variables listed in Table 4.1. Means for each quartile along the horizontal axis (0-25%, 25-50%, etc.) are shaded in gray, as a visual guide to the general upward trend but also the general non-linear shapes of these relationship.

Figure 4.1: Dependence of Section Means for Student Gains on Observed Student-Centered Class Time



Overall, the more class time spent on student-centered activities, the higher the student gains, although gains appeared to plateau slightly at the high end. This apparent plateau is intriguing. We expect student-centered activities to facilitate learning, but some amount of framing and signposting by the instructor is also vital to shape and guide student learning and help students recognize and relate the ideas that emerge from their work.

4.3.2 Learning gains of math-track students and pre-service teachers related differently to the total student-centered time.

The relationship between student gains and the composite classroom variables was further complicated by the two types of students represented in the data set: math track students and pre-service teachers. As discussed in Chapters 3 and 5, the two student groups were quite different. Pre-service teachers started out with less interest in and weaker enthusiasm for mathematics (see Section 3.4.3) and reported lower gains overall (see 3.2.1). Correlation analysis and inspection of scatter plots of the bivariate data revealed that the nature of the relationship between learning gains and total student-centered time was somewhat different for math-track and pre-service course sections. Briefly, the response of math-track students appeared to be less linear than that of pre-service teachers. These differences may indicate a difference in the balance of student-centered activities and instructor scaffolding that benefited these groups most, but the numbers of sections of each type is too small to demonstrate this definitively.

Such differences in student groups also complicated the overall relationship between student gains and composite classroom variables. Dividing the data set yielded two rather small sub-group sets: 25 math track and 11 pre-service sections. This reduced statistical power and constrained our ability to deeply understand relationships between the variables. The statistical technique called Hierarchical Linear Modeling (HLM) discussed in Section 4.5 allowed us to control for those differences without sacrificing statistical power.

4.4 Relationship between Student Gains and Observer Survey Variables

Observers' ratings of classroom interactions and atmosphere, discussed in detail in Chapter 2, are also of interest to this analysis. The 14 indicators fell into four distinct clusters with high internal reliability (see Table 2.4). Based on these clusters, we constructed four composite variables measuring extent and quality of *student-instructor interactions*, *student-student interactions*, *students' role in setting course pace and direction,* and *instructor behaviors.* These composite variables are related not just to an instructor's choice of instructional activities, but to the quality of interpersonal interaction among students and instructors, the degree to which students share responsibility for the success of the course, and instructors' efforts to set a positive tone and provide appropriate guidance. Therefore these should reflect different aspects of the course than do the classroom time variables.

4.4.1 The relationships between observer ratings variables and student gains were mostly strong and linear.

Three of four observer survey composite variables correlated well with most of the student learning gains variables. The relationships between select student gains, averaged for each course section, and the observer survey variables are summarized in Table 4.2.

The correlations between student gains and student-instructor and student-student interactions were rather high and statistically significant. The correlations between students' role in setting course pace and direction and learning gains were even stronger, while still statistically significant. Thus, these classroom traits appear to have directly affected students' learning gains: students reported greater learning when these factors were more strongly present.

Table 4.2: Correlations of Section Means for Select Student Gains and Observer Ratings Variables

Student Gains	Student-instructor interactions	Student-student interactions	Students' role in setting course pace & direction	Instructor behaviors
Mathematical thinking	0.396*	0.424**	0.633***	0.036
Affective gains	0.427**	0.492**	0.542**	0.062

On the other hand, instructor behaviors showed no relationship to student gains. This composite variable consists of two ratings: the extent to which instructors established a positive atmosphere and how well they summarized or placed class work in a broader context. In Chapter 2 (Section 2.2.7) we reported that observers rated most instructors highly on those items. That is, most instructors attempted to create a positive atmosphere and to scaffold the material, whether or not they chose student-centered instructional approaches. Thus it was not primarily instructors' good intentions that yielded positive outcomes for students—these may be necessary but not sufficient traits, that do not appear to be rare. Rather, the quality of student-instructor and student-student interactions, and especially students' shared responsibility for setting class pace and direction, made a measurable difference in the strength of student learning gains.

4.5 Hierarchical Linear Modeling (HLM): How Student Gains Depend on Course Practices

The approach used so far, comparing section-averaged gains with section-level observation data has certain down sides. As each point in the data set represents an aggregated response of many students, it essentially collapses many students' varying experiences into a one point—yet clearly every student in a course does not respond identically to instruction. Thus, this approach obscures some of the variability of student experience. The statistical tool of Hierarchical Linear Modeling (HLM) addresses this limitation by simultaneously accounting for both class-level variables, such as total student-centered time, and for student gains on an individual basis, without aggregating them and thus artificially reducing their inherent variability. In a modeling approach, we do not just look for associations, but use them to make a prediction for how student outcomes relate to the combined effects of multiple factors.

4.5.1 Participating in an IBL course was a predictor for positive student outcomes.

We used a Hierarchical Linear model to assess the relative impact of participation in the IBL program on self-reported student gains. The model incorporates the influence of selected student characteristics (gender, class year) and course characteristics (including participation in the IBL program) to predict student gains from the course. Thus, Hierarchical Linear Model provides a more precise and balanced means of assessing the relationship between class type and learning gains by accounting for all types of variability due to differences in individual students and the effects of class membership.

The outcome variable for the model is a composite variable of weighted factor scores for five learning gains survey items: gains in the main concepts explored in class, the relationships among the main concepts, students' own ways of mathematical thinking, how mathematicians think and work, and how ideas from this class relate to ideas outside mathematics. The composite variable showed high internal reliability at $\alpha = 0.88$.

The first model examined 1239 student survey responses from 80 course sections. The model compared the outcome variable for IBL and pre-service courses (course-level or level-2 variables), gender and class level (student-level or level-1 variables) using a two-level hierarchical model. For the binary variables, IBL and pre-service course were coded 1 for participating, 0 for not; gender and class level were treated similarly. Table 4.3 summarizes the outcomes of the two-level model.

Table 4.3: Two-level Hierarchical Linear Model for Full Survey Data Set

Independent Variables	Coefficient	Standard Error	Significance
Intercept	-0.22	0.10	0.030
Course-level variables (77 degrees of freedom)			
IBL	0.33	0.09	0.001**
Pre-service	-0.48	0.17	0.007**
Student-level variables (1214 degrees of freedom)			
Gender	0.05	0.06	0.367
First-year	0.16	0.07	0.026*
Sophomore	0.15	0.08	0.070*
Junior	0.02	0.09	0.815
Senior	-0.01	0.08	0.870

As shown in Table 4.3, participation in an IBL course section had a positive (0.33), strongly statistically significant effect on the composite learning gain variable. Thus, even after accounting for all the course and student-level variation, taking an IBL course predicted higher student learning gains than did taking a non-IBL version of the course. This confirms the results found earlier (see Section 3.2.1) and, as we will show, is corroborated by findings on subsequent student grades presented in Chapter 6 (Section 6.2.1).

4.5.2 Participating in a course for pre-service teachers was a negative predictor of student learning, showing that these students have distinct mathematics learning needs.

The Hierarchical Linear Model also showed that enrollment in a pre-service teacher course, however, had a negative (-0.48) and strongly statistically significant effect on the composite learning gain variable. That is, pre-service teachers are predicted to report lower learning gains than math-track students, even after controlling for other course-level and student-level variations. Once again, this confirms our findings (Section 3.2.1) that pre-service teachers respond differently to college mathematics courses than do math-track students, reflecting the different needs, interests, and expectations of this population. In this study, courses targeted to pre-service teachers were also IBL courses—so the positive effect of IBL essentially counteracted the mathematics learning challenges posed by this student group. It is unfortunate for the study that no non-IBL pre-service courses could be included—but perhaps very fortunate for students themselves, as this model would predict rather poor results. We suggest that pre-service teachers be viewed as "canaries in the coal mine"—a group that may be particularly sensitive to poor mathematics teaching, but also particularly responsive to improvements therein.

4.5.3 Some student-level variables were significant predictors of student learning.

Variables indicating students' status as college first-years (0.16) or sophomores (0.15) had positive and statistically significant relationships to student gains. That is, students early in their college careers tended to make higher learning gains than did those farther along. This finding sheds some light on our results on benefits of IBL to younger students (Section 3.2.4)—they may be generally responsive to new college experiences, as well as especially responsive to IBL.

No other student-level factors showed a statistically significant effect on the student learning gains in the model. It seems particularly puzzling that gender does not appear as a significant predictor in the model, since gender differences were strong in other statistical tests. We continue to investigate other student-level factors that may relate to gender.

4.5.4 The observed total percentage of class time spent on student-centered activities was the strongest predictor of student outcomes.

The strong effect of the "IBL" variable in the HLM shows that designation of a course as "IBL" or not was meaningfully related to student outcomes. However, this campus-assigned designation is a simple, binary variable that does not capture the real variation in instructional practices in the course (e.g. Figure 2.2). Thus, we explored the direct connection between student learning gains and observation variables already shown (4.3.1) to reflect the use of student-centered instructional activities. Adding observation variables to the model discussed above would be the best test from this perspective. But the observation and survey samples were not exactly matched (4.1). Therefore, including the observation variables in that model would lower the sample size and reduce statistical power.

Instead we constructed another model, where that smaller sample would still suffice. This model examines the effect of three course-level variables on student outcomes: total percentage of the student-centered time, plus the previously tested indicators of IBL and pre-service courses. This model does not use any student-level variables (such as gender and class year) as predictors of outcomes. The model relies on data from the 36 course sections with 670 students where observations were conducted. Table 4.4 summarizes the results from this model.

Table 4.4: Two-level Hierarchical Linear Model for Combined Observation-Survey Data Set

	Coefficient	**Standard Error**	**Significance**
Intercept	-0.32	0.104	0.004**
Course-Level Variables (32 degrees of freedom)			
IBL	-0.24	0.240	0.335
Pre-service	-0.35	0.130	0.010*
Total % class time spent on student-centered activities	0.01	0.004	0.014*

As Table 4.4 shows, the model predicted positive and significant effects for student-centered activities, negative and significant effect for pre-service variable, and no significant effect for IBL. That is, an increase in total student-centered time had a positive impact on the student learning gains. But when this was included in the model, IBL lost significance as a predictor.

Since the IBL effect, which was very strong in the previous model, disappeared after adding the observation-based variable, it is apparent that the real distinction underlying the IBL/non-IBL label is the emphasis on student-centered activities during class time. Indeed, the essence of IBL is its focus on student-centered teaching and learning approaches, and the total percentage of student-centered time variable captures this essence better than does the institutional designation of IBL (see Section 2.3).

4.6 Conclusion, Strengths and Limitations of this Analysis

Overall, several lines of evidence indicate that good student outcomes—as measured by our most broadly applied learning indicator, the student surveys—result from the use of student-centered active teaching and learning approaches. Correlations reveal positive though non-linear relationships between student gains and the observed percentage of class time spent on student-centered activities or with students in the lead role, and with observer ratings of interpersonal interactions and shared responsibility. High values of these were common features of the IBL courses in this study.

Hierarchical Linear Modeling is a more sophisticated approach that honors the nested structure of the data, helping to avoid estimation bias and "false positive" errors that would lead to over-stated claims. HLM analyses thus allow us to estimate the effects on learning of variables at both the course and student levels. The HLM results show that the observed percentage of class time spent on student-centered instructional activities was a quite strong and positive predictor of student learning. HLM also predicted better outcomes for first-year and sophomore students, similar to other findings we report.

Pre-service courses, however, were negatively associated with learning, suggesting that these students are not so easily reached by college mathematics learning experiences as are their math-track peers. Yet these students will play crucial roles in shaping the mathematical knowledge and learning attitudes of the next generation (Beilock, Gunderson, Ramirez & Levine, 2010). Thus it is particularly important that they respond well to the small number of mathematics courses that they will take in college. Overall, the hierarchical model suggests that the good effect of IBL can counter what is in essence a negative starting line for IBL teachers. This may be taken as validation for the choice to focus IBL teaching resources on this group of students.

Where observation data was not available, institutional designations of IBL or non-IBL courses proved to be a useful substitute. The positive linkage between the IBL label and the student outcomes shows that, at these sites, the IBL-labeled courses did in fact use student-active approaches that fostered learning. That is, a dose of “IBL” had a positive effect; but a dose of student-centered activity had a stronger and more easily measured effect. Ultimately, researchers should rely on ground-truthed data from classroom observation to document actual instructional practices, independent of semantic labels. Yet direct observation poses challenges of time, cost, observer experience or training, and inter-rater reliability. Seeking to balance these factors, our observation protocol clearly captured some important aspects of instruction, but likely missed other relevant details.

This type of analysis is not even possible without a very large volume of data, and the strength of the relationships uncovered so far is encouraging. However, even with over 1200 student survey responses and 300 hours of observation, the size of the data set is on the very edge of what is needed to extract good correlations or to construct a complete, two-level hierarchical model. The cost and effort required to document these linkages may help to explain the paucity of well-controlled studies that link student outcomes to innovative learning approaches in STEM disciplines (Hough, 2010a,b; Ruiz-Primo, Briggs, Iverson, Talbot & Shepard, 2011).

Some puzzles remain for future analysis. Relative to other chapters, these findings are more preliminary—not final answers to our research question, but tantalizing hints about a story that is not yet complete.

4.7 References Cited

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers’ math anxiety affects girls’ math achievement. *Proceedings of the National Academy of Sciences*, *107*(5), 1860-1863.

Hough, S. (2010a). The effects of the use of inquiry-based learning in undergraduate mathematics on student outcomes: A review of the literature. [Report prepared for Ethnography & Evaluation Research, University of Colorado at Boulder.]

Hough, S. (2010b). The effects of the use of inquiry-based learning in undergraduate mathematics on prospective teacher outcomes: A review of the literature. [Report prepared for Ethnography & Evaluation Research, University of Colorado at Boulder.]

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*(3), 223-231.

Ruiz-Primo, M. A., Briggs, D., Iverson, H., Talbot, R., & Shepard, L. A. (2011). Impact of undergraduate science course innovations on learning. *Science*, *331*, 1269-1270.

Chapter 5: Findings from Tests of Mathematical Knowledge and Thinking

5.1 Introduction

In addition to asking students about their learning, we wished to directly study changes in students' mathematical content knowledge. After searching for suitable tests and discussing test opportunities with instructors and others, we chose two different instruments for this purpose. The first test, called *Learning Mathematics for Teaching* (LMT), offered us well-validated instruments for measuring pre-service teachers' cognitive gains from an IBL mathematics course. The other test, which we called the *Proof Test*, measured students' ability to evaluate a mathematical argument and determine its validity. Because of the great variety in courses taught at the IBL Centers, each test was appropriate for students from a subset of those courses.

Using these two tests, we gathered two modest data sets on the development of students' mathematical knowledge and thinking. These studies addressed our research questions:

- How do students' mathematical knowledge and thinking change during an IBL college mathematics course?
- How do the changes differ by student groups, especially between IBL and non-IBL students?
- How do the changes align with results from other measures of students' learning gains?

Below we first report findings from the LMT test given to pre-service teachers, then results from the proof test given to a set of math-track students.

5.2 Changes in Pre-Service Teachers' Mathematical Knowledge for Teaching

5.2.1 The Learning Mathematics for Teaching (LMT) instrument and study samples

The LMT instruments were developed and validated by a team at the School of Education, University of Michigan, for assessing professional development courses for K-12 mathematics teachers. The items are designed to measure the development of mathematical knowledge needed for teaching: solving problems, using definitions, and identifying adequate explanations (Hill, Schilling & Ball, 2004). Items on the LMT instrument for *Number Concepts and Operations* reflect both the content that teachers teach and the special knowledge they need to teach that content to students. Detailed description of our methods is provided in Appendix A5.1.

A total of 109 pre-service teachers took both the pre-test and the post-test at two campuses in either of two academic years. We collected data from three different groups of students, each of whom participated in a separate two-course IBL mathematics sequence. None of these courses had comparative (non-IBL) sections. Table 5.1 summarizes the study sample—which includes essentially all students enrolled in these courses—and shows its skewed gender distribution.

Table 5.1: LMT Test Sample of IBL Pre-Service Teachers, by Course Group and Gender

Study Group, by Course		**Women**	**Men**	**Total**
Group 1	*Elementary grades*	27	0	27
Group 2	*Elementary & middle grades*	64	2	64
Group 3	*Secondary grades*	9	9	18
Total		100	11	109

The pre-test consisted of 24 multiple-choice items given at the start of the two-course sequence, and the post-test of 23 matched, but not identical, items given at the end. Raw scores from the pre- and post-tests were converted into standardized IRT (item response theory) scores using scoring tables provided by the LMT developers. Use of the standardized IRT scores enabled us to equate scores on the pre-and post-test, and ensured reliability of the results. In analyzing and reporting the data, we use these standardized IRT scores.

5.2.2 Students made clear progress in learning mathematical knowledge for teaching.

The three groups started at different initial levels of knowledge for teaching elementary number concepts and operations. On average, Group 3 outperformed Group 1 on the pre-test ($p<0.05$). This may be due to the higher number of prior college mathematics courses taken by Group 3 students—typically six or more—as compared to the other groups ($p<0.001$).

All three groups made statistically significant score increases from pre- to post-test. Table 5.2 displays these increases. On average, students in Groups 1 and 2 gave two additional correct answers on the post-test, while Group 3 students improved by three additional correct answers. That is, students with a stronger math background improved their score more than students with less previous college math experience. The increases showed effect sizes (above 0.70) between the pre- and post-test for Groups 2 and 3 that are generally considered to be large (Cohen, 1988). The increase for Group 1 was nearly as large.

Table 5.2: Average Changes in Scores for a 24-item Test, by Course Group

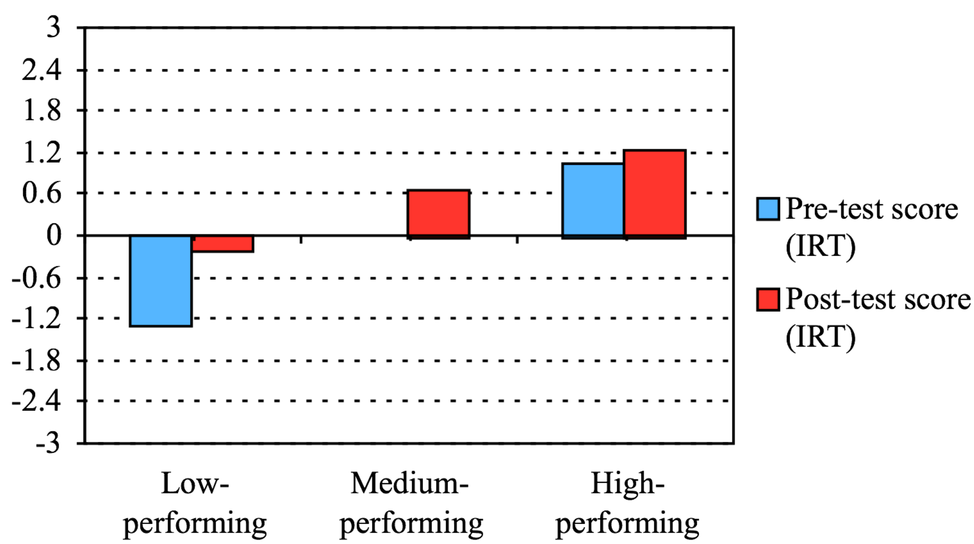
Student Group	**Change in Average Score**	**Sig. Level**	**Effect Size**
Group 1 (*N=27*)	+2	$p<0.01$	0.67
Group 2 (*N=64*)	+2	$p<0.001$	0.79
Group 3 (*N=18*)	+3	$p<0.01$	0.90

Pre-service teachers clearly made progress in their IBL classes in learning the mathematical knowledge needed for teaching K-12 mathematics. The extent of learning did not depend on the level of teaching certificate they were pursuing. Previous work has shown that teachers' increased ability to understand and communicate the needed mathematical definitions (as measured by the LMT items) has a positive effect on their actual classroom instruction (Hill, Schilling & Ball, 2004). Therefore these test score gains imply that these pre-service teachers are being prepared well for their future classroom duties. Although the LMT instruments have previously been used with in-service teachers, our results indicate that these instruments are suitable for measuring MKT among pre-service teachers as well.

5.2.3 Test score improvement was strongest among weaker students.

We explored possible variation in LMT test results with respect to students' initial test performance. Correlation between the initial test score and test score gain was clearly negative ($r= -0.362$, $p<0.01$), meaning that students with initial lower scores had higher test score gains than students who performed better at the beginning of a course. We studied this result more closely by dividing the students into three performance groups (low, medium, high) on the basis of their pre-test score: 50% or fewer correct answers, 51-75% correct, and more than 75% correct. Figure 5.1 shows the average change in standardized test score for the three groups.

Figure 5.1: Changes in the Standardized IRT Test Score, by Performance Group
(initial score ≤ 50%: N=22; 51-75%: N=60; ≥ 75%; N=27)



This analysis reveals interesting differences in the test score gains for students who start with weaker or stronger initial scores. Low- ($p< 0.01$) and medium-performing ($p< 0.05$) students improved more than did high-performing students. On average, initially low-performing students provided 4-5 additional correct answers on the post-test, compared with three additional correct answers by initially medium-performing students, and no more than one additional correct answer from initially high-performing students.[1] Weak students clearly benefited strongly from

[1] This is not a ceiling effect: as a group, high-scoring students had room for further score improvement.

their IBL mathematics course. This result is confirmed by further findings reported in Section 5.2.5.

5.2.4 Other group differences were minor.

We compared pre- to post-test changes in LMT test scores among various sub-groups but found no substantial differences.

- Most students were juniors (64) or seniors (37). We saw no difference in score changes related to students' academic status.
- The number of men was small. We found no difference in women's (98) and men's (11) score changes.
- We found no difference between students who reported no teaching experience (87) and those who had at least some prior experience in teaching (22).
- We classified students into three groups based on the number of previous college mathematics courses they had taken: 0-1, 2-3, 4 or more. Students who had taken four or more mathematics courses outscored the two other groups on both the pre- ($p<0.05$) and post-test ($p<0.001$), but their scores did not improve more from pre- to post-test.
- Test score changes were similar for Hispanic or Latino versus non-Hispanic/non-Latino students.
- Most of the students were white (75). No differences in LMT score gains were observed upon comparing white students, Asian students, and students of other races.
- We found no differences in test score changes by major, comparing mathematics/applied mathematics majors, science majors, and non-science majors.

These results indicate that pre-service teachers' learning of elementary MKT from an IBL course did not clearly depend on their academic status, teaching experience, their experience of college mathematics, academic major, or on demographic characteristics. That is, pre-service teachers with varying academic goals and backgrounds benefited equally from an IBL course targeted to learning mathematics for K-12 teaching.

5.2.5 Test score improvement was not related to other learning indicators.

At the beginning of their IBL courses, students reported their score (1-5) on the AP Calculus test (if they had taken it), their current estimated undergraduate GPA, and their expected grade in the present course. They also reported an expected course grade at the end of the two-term sequence. Finally, we conducted a small experiment in which we asked instructors to separately rate both their students' mathematical expertise and students' learning from the course. We checked to see how well these other academic indicators correlated with students' LMT scores. Detailed results of these analyses are provided in Appendix A5. Briefly, they lead us to conclude that:

- the LMT test did measure aspects of knowledge that are mathematical in nature, even though it is distinct in its focus on mathematics for teaching

- students who expected to get a good grade did get better LMT test scores, but the expected grade was not well linked to students' test score *growth*
- instructors were more accurate in assessing students' mathematical expertise than in assessing student *learning* as judged by test score growth.

In other words, the results suggest that both students and instructors judged student performance relative to other students with some accuracy, but did not accurately predict learning or growth.

5.3 Changes in Students' Ability to Evaluate Mathematical Arguments

We used another test, nicknamed the proof test, to measure different aspects of students' mathematical knowledge and thinking. This test studied students' ability to evaluate a mathematical argument and determine its validity. Below we report findings from the proof test.

5.3.1 Proof test instrument and study samples

The proof test was based on assessment items on evaluating mathematical arguments that were designed by Weber (2009). Since critiquing proofs is an important aspect of many IBL courses (and comparable non-IBL courses) and a skill that can be tested across a variety of topical courses, we set up test situations on two different campuses. The first data set was gathered from one-on-one problem-solving interviews. Later, the test was revised into a paper-and-pencil form that was administered either in class to all students, or out of class to volunteers.

In the interviews, students were asked to verbally explain the reasoning behind their answer about each argument, and these responses were recorded. On the paper-and-pencil test, students wrote down their reasoning about whether or not each argument was a mathematical proof. The data came from 42 IBL students (27 men, 15 women) and 35 non-IBL students (19 men, 16 women) at the end of a mathematics course that was, according to instructors, an introductory or mid-level proof-based course (although in practice we found that many students had substantial proving experience in prior courses). Most of the students were volunteers (63) who were paid a modest honorarium for participating. Only 14 students took an in-class post-test. In addition, we got pre/post-test data from one section (20 pre-, 14 post-tests).

The one-hour test included nine arguments on algebra, number theory and calculus. Each argument was followed by structured questions to probe:

- Did students understand the argument?
- To what extent did they find it to have explanatory power?
- To what extent were students convinced by the argument?
- Did students consider the argument to be a mathematical proof?

After answering these questions, students reported their overall view of why they did or did not consider each argument to be a mathematical proof. Details about the test, data and analysis methods are given in Appendix A5.

5.3.2 Students did well on the proof test.

Table 5.3 summarizes IBL and non-IBL student responses for the three valid and the six invalid arguments. (In addition, Appendix A5 includes results for individual problems.)

Table 5.3: Average Student Ratings of Nine Mathematical Arguments

Criteria for evaluating each argument	**Rating, scale 1 to 5**		Interpretive comments
	IBL	Non-IBL	
Understanding	*1=do not understand fundamental details to 5=understand completely*		
Valid arguments	4.6	4.5	Both groups felt they understood the arguments very well.
Invalid arguments	4.4	4.5	
Conviction	*1=not convinced at all to 5=completely convinced*		
Valid arguments	4.5	4.3	Students were less convinced by invalid than valid arguments, but they found invalid arguments fairly persuasive.
Invalid arguments	2.9	2.8	
Explanatory power	*1=does not explain to 5=really illuminates why it is true*		
Valid arguments	4.3	4.3	Students rated explanatory power with less clarity than the other criteria. They attached rather strong explanatory power to both invalid and valid arguments.
Invalid arguments	3.1	3.0	

Students reported little difficulty in understanding any of the arguments. They felt they understood the arguments well and appeared to be influenced by both the persuasiveness and explanatory power of the proofs in assessing argument validity. In general, students were more convinced by valid arguments than by invalid arguments, but often found the invalid arguments fairly convincing. Students also perceived more explanatory power in valid than in invalid arguments, though they also attached fairly strong explanatory power to invalid arguments.

On the whole, both IBL and non-IBL groups were fairly successful in distinguishing valid from invalid proofs. Most students considered the valid arguments to be mathematical proofs: 86 to 95% of IBL students and 76 to 91% of non-IBL students, depending on the argument. However, students were somewhat less successful in identifying invalid arguments. Depending on the argument, 17-100% of IBL students and 20-86% of non-IBL students considered invalid arguments not to be mathematical proofs. Invalid arguments 2, 8 and 9 were particularly challenging for students to identify (see Section 5.3.3). Overall, students' ability to discriminate arguments approached that of experts, but was not yet fully developed (Hersh, 1993; Harel & Sowder, 1998a, 1998b; Selden & Selden, 2003).

5.3.3 IBL students' overall scores only slightly exceeded those of non-IBL students.

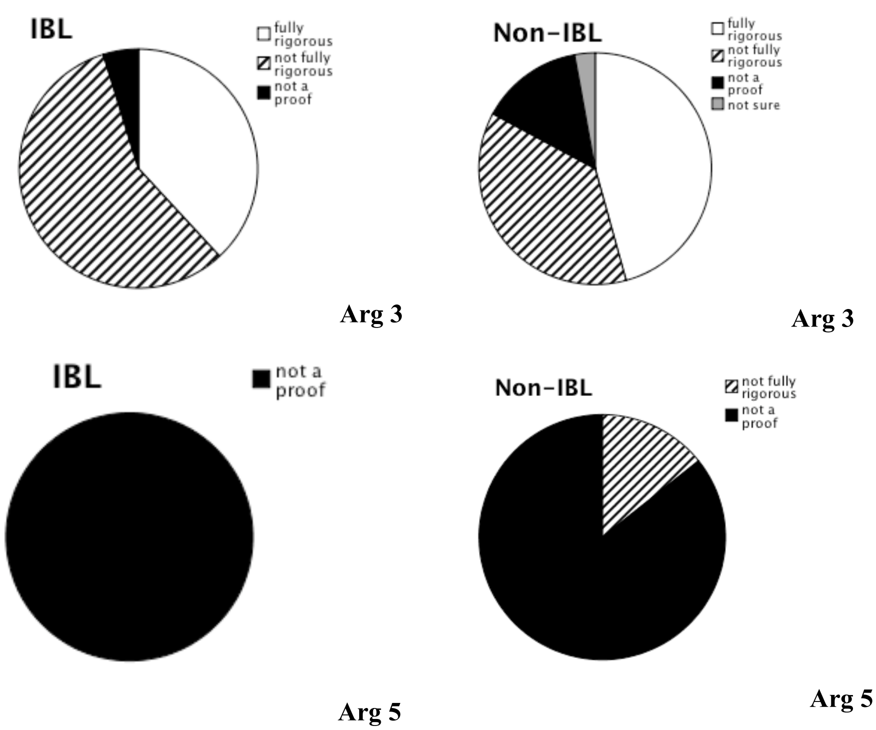
Overall, the quantitative data revealed few clear differences between IBL and non-IBL students' assessment of these arguments. This is probably explained by the fact that most of the students were volunteers and were strong students, reporting an A or B grade-point average: all of the students were fairly competent in assessing the arguments.

A few minor differences did appear in these data. No difference between IBL and non-IBL students were seen in the overall average scores for all nine arguments on any criteria: understanding, conviction or explanatory power. However, IBL women generally understood the arguments slightly better ($p<0.05$) than did non-IBL women. Moreover, qualitative analysis of students' reasoning displayed some differences in favor of the quality of IBL students' thinking (see Section 5.3.6).

IBL students succeeded slightly better in recognizing valid arguments than did their non-IBL peers. In all, only 2% of IBL students, but 14% of non-IBL students, erroneously labeled two out of three valid arguments as not qualifying as proofs. Additionally, 24% of IBL students (10) but just 17% of non-IBL students (6) were able to rule out all six invalid arguments as fully rigorous proofs. That is, IBL students were somewhat less likely to erroneously identify a valid argument as invalid and, in turn, also less likely to identify an invalid argument as a fully rigorous proof. Generally, IBL students were slightly more critical in judging a proof to be fully rigorous. This applied to valid arguments 1 and 3, and to invalid arguments 6-8.

Individual problems showed some interesting differences between IBL and non-IBL students (Figure 5.2), but these differences were not consistent across the problems.

Figure 5.2: Examples of the Distribution of Students' Answers to Specific Problems



Valid Argument – deductive number theory

IBL students are more critical, more often seeing this argument as not a fully rigorous proof.

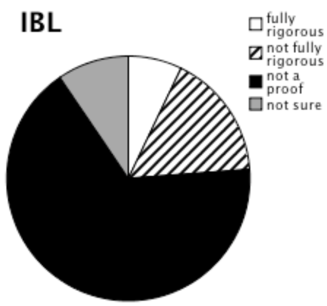
Non-IBL students more often see this argument, erroneously, as not a proof at all.

Invalid Argument – empirical number theory

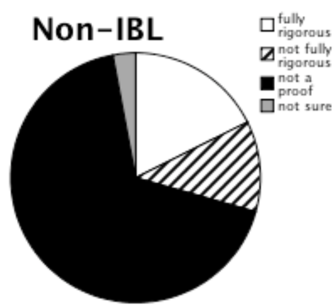
All IBL students see this argument as not a proof.

Some of the non-IBL students see this argument, incorrectly, as a proof.

The difference is statistically significant ($p<0.05$).



Arg 7

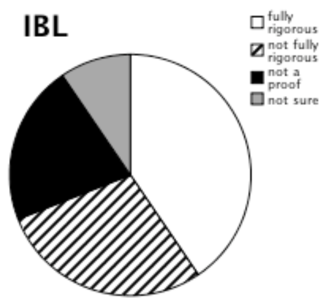


Arg 7

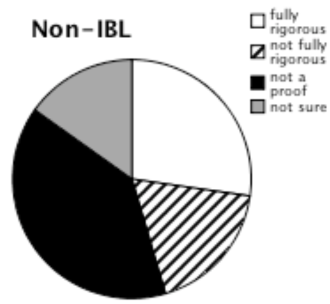
Invalid Argument – deductive number theory, invalid proof structure

Both groups largely see this argument, correctly, as not a proof.

Non-IBL students more often identify this argument erroneously as a fully rigorous proof.



Arg 9



Arg 9

Invalid Argument – deductive algebra, invalid warrant

IBL students more often are fooled by this argument and see it as a proof.

Non-IBL students more often see this argument correctly as not a proof, but with some hesitancy.

In only one case (Argument 5, invalid empirical number theory) did we observe a statistically significant difference between IBL and non-IBL students' validation of arguments ($p<0.05$). All the IBL students (100%) but fewer non-IBL students (86%) correctly saw that it didn't meet the standards of a proof. Easily half of both IBL (55%) and non-IBL (43%) students recognized that the perceptual calculus argument (6) was not a proof. Most (67% IBL, 66% non-IBL) also recognized the invalid proof structure in Argument 7. However, non-IBL students slightly more often identified this argument as a fully rigorous proof. On the other hand, nearly all students were fooled by the invalid Argument 2 (perceptual algebra). Only 17% of IBL and 20% of non-IBL students labeled this as not a proof.

Arguments 8 and 9 were also more difficult for students. In all, 26% of IBL and 31% of non-IBL students considered invalid Argument 8 (deductive calculus) to be a fully rigorous proof. While IBL students were more often on the right track in assessing the arguments, an exemption applied to invalid Argument 9 (deductive algebra). IBL students more often incorrectly viewed this as a proof: 41% of IBL but 26% of non-IBL students labeled the argument as fully rigorous.

Overall, results from the quantitative proof test data indicated that IBL students did only slightly better than non-IBL students in assessing arguments of the proof test after a mathematics course.

5.3.4 Assessment of arguments did not depend on students' personal or academic background.

The circumstances for proof test interviews was different than for the paper-and-pencil test, and some differences did appear in students' performance for these two cases. Students who were interviewed understood the arguments less clearly ($p<0.01$), and could slightly less often ($p<0.05$) provide a good explanation for the valid arguments. This result may be due to the higher-stress testing situation of what is essentially an oral exam.

On the other hand, the interviewed students did better at analyzing two of the invalid arguments. They were more certain (71%, $p<0.05$) than paper-tested students (40%) that the invalid perceptual calculus argument was not a proof. Moreover, 42% of the paper-and-pencil students but only 18% of the interviewees considered the invalid algebra Argument 9 as a fully rigorous proof ($p<0.05$). In sum, the type of test (paper-and-pencil vs. interview) had little influence on students' performance. Paper tests had a clear advantage in increasing the amount of data that could be acquired for a given investment of researcher time.

We explored possible differences in performance between various sub-groups but found no clear differences. We found no general gender difference in validation of the arguments. However:

- Women (40%) more often than men (23%) incorrectly considered the invalid calculus Argument 8 to be a fully rigorous proof.
- Comparison of IBL and non-IBL women, and separately of IBL and non-IBL men, revealed no differences.
- No difference was found by students' academic status.
- No difference was found by student race/ethnicity.

Overall, performance in the assessment of mathematical arguments did not depend on personal or academic background.

5.3.5 IBL students used more advanced criteria in reasoning about mathematical arguments.

We examined students' written reasoning for each argument. These qualitative data were coded and analyzed qualitatively according to eight main themes related to the quality of students' assessment of the arguments (see Appendix A5):

- Understanding
- False statements
- Inadequate reasoning
- Use of steps as a criterion
- Formalism
- Quality of explanation or reasoning
- Rigorousness
- Assessment of an argument as a whole.

These themes were used to code the type of reasoning students gave when assessing the argument, and the features of the mathematical arguments upon which they commented. Results are based on written comments by 54 students (28 IBL, 25 non-IBL). Appendix A5 includes the thirty sub-categories under the main themes and the frequencies in each category. The categories were derived from preliminary analysis of a subset of written comments and finalized during further analysis of the complete data set.

Compared to IBL students, non-IBL students more often used *inadequate* reasoning ($p< 0.05$). This category included lack of justification (e.g., "I feel this is a proof"), use of perceptual or empirical evidence, and considering an invalid argument as complete. In all, 46% of non-IBL but only 24% of IBL students provided an inadequate assessment at least twice. Moreover, 54% of non-IBL students' assessments lacked a justification, versus 40% of IBL students' assessments. Non-IBL students also more often ($p< 0.01$) made *false* statements (e.g., "Lines 1-7 have nothing to do with true claims," "Don't take into account negative numbers"). This happened especially in their reasoning about invalid arguments. In all, 40% of non-IBL but just 14% of IBL students made at least one false statement about an argument. These results suggest that non-IBL students used less sound reasoning in assessing the arguments.

In other categories, IBL and non-IBL students performed similarly. For example, both IBL (52%) and non-IBL (51%) students detected a flaw in an invalid argument at least once. Their assessments also were equally based on *formalism*, as they noticed the inclusion or absence of theorems, mathematical rules, or formal structure in the arguments (IBL 50%, non-IBL 54%). They made statements such as "Uses fundamental algebra without adding theorems." or "Logical structure is valid but not written formally." Emphasis on form is suggested by Sowder and Harel (2003) to represent a ritual and external conviction proof scheme. Students' use of steps as a criterion for assessing arguments also focused more on the external structure of arguments. This criterion contrasted with a more sophisticated emphasis on meaning, internal logic and reasoning (cf., Smith, 2006). Non-IBL students only slightly more often than IBL students based their reasoning on the *use of steps* in arguments (IBL 31%, non-IBL 43%), making statements such as "The steps were acceptable" or "Proved step by step."

Students' understanding of an argument was used as a criterion for judgment by 31% of IBL and 40% of non-IBL students. Another criterion in students' comments reflected an overall outlook on arguments. Instead of pinpointing single steps or notes in an argument, students assessed general style, beauty, and the topic or mathematical level of an argument. No general difference appeared in these judgments: 33% of IBL and 31% of non-IBL students discussed these characteristics at least three times. However, compared with non-IBL students (40%), IBL students (48%) slightly more often assessed the way of writing or presenting an argument. The style in which proofs are written is important in validating proofs also among mathematicians (Selden & Selden, 1999). In contrast, non-IBL (31%) more than IBL (14%) students made general statements about the outlook (beauty, easiness) of an argument.

Compared with non-IBL students, IBL students more often based their assessment on the quality of *explanation or reasoning* in an argument ($p< 0.05$). This occurred more often with valid than invalid arguments. In all, 48% of IBL students, but only 20% of non-IBL students, wrote about the level of explanation and reasoning three times or more. In contrast, 74% of non-IBL but just 48% of IBL students did not seek further explanation or reasoning in valid arguments. Moreover, IBL students slightly more often pointed to *rigor* of the arguments. Some students outlined a more rigorous proof of an argument, and IBL students did this more often (19%, vs. 9% of non-IBL students). However, these differences were not statistically significant. Overall, IBL students were more critical in assessing the arguments. Their emphasis on rigor together with explanatory power of arguments and construction of sub-proofs in assessing arguments suggests a level of sophistication in validating mathematical proofs (Hanna, 2000; Hersh, 1993; Selden & Selden, 2003; Pfeiffer, 2009; Weber, 2008).

These results indicate that, while IBL students were slightly stronger than non-IBL students in assessing the arguments, their reasoning was also partly based on rules and formal structures of mathematical arguments. However, the differences in student language and reasoning suggest that IBL students were using different criteria to judge arguments—criteria that may be similar to those used by expert mathematicians in viewing proofs as rigorously constructed, explanatory arguments.

5.3.6 Modest positive differences were seen between students tested before an IBL course, and those tested afterwards.

Because we wanted to examine students' development throughout a course in assessing mathematical arguments, we designed a separate experiment in collaboration with one IBL instructor. He gave the proof test to his class at the end of one course and again at the start of a different section of the same course the next semester. A true pre/post comparison with the same students taking both tests was not available, but the instructor saw no *a priori* reason that the student groups in the two terms would not be comparable. Results from these two data sets were then compared.

A clear positive difference ($p<0.05$) in students' ability to recognize valid proofs was seen for only one argument. Among the pre-tested group, 35% of students labeled invalid Argument 7 (invalid proof structure) as a fully rigorous proof, but among the post-tested students, none did so. They also ascribed more explanatory power to the valid arguments at the end of the course ($p<0.01$). On the other hand, post-tested students were more convinced by both the three valid ($p<0.01$) and the six invalid ($p<0.05$) arguments at the end of the course.

A few differences in written assessments appeared between pre- and post-tested students. In all, 50% of the pre-tested students gave at least three inadequate statements, whereas post-tested students gave no more than 1-2 inadequate statements ($p<0.01$). Moreover, 57% of post-tested students requested more explanation for valid arguments, whereas only 20% of pre-tested students did so ($p<0.05$). Post-tested students were also slightly more critical about invalid arguments: 57% of them but only 15% of pre-tested students offered critiques at least five times.

Post-tested students also tended to assess the general outlook of an argument more than did pre-tested students, especially for invalid arguments ($p<0.05$).

Assuming, as the instructor did, that the student populations in the two sections were comparable, these findings suggest that the proof test was a moderate indicator of students' learning to assess mathematical arguments. It can be argued that students learned to make more adequate statements about mathematical arguments and looked at them more critically after an IBL course, although they were still easily convinced by both valid and invalid arguments. However, this experiment did not provide a true pre-/post-test setting, and the number of the students in the two sections was very small. They also were good students, as reflected in their expected grade from the course. We are unable to compare these results to those from a more traditional proof-based course.

5.4 Discussion: Strengths and Limitations of the Test Data

We met with modest success in our efforts to gather test data directly measuring mathematical content knowledge across IBL and non-IBL sections at the campuses. Ideally, we would administer a single test to many different students across multiple courses or at least multiple sections, but no such test was available. Taking another approach, we tried to engage instructors to develop and administer common test items in IBL and comparable non-IBL sections. We were not successful at any campus where this opportunity existed.

Instead, we used two different tests for separate groups of students, each given at two campuses. These tests measured students' mathematical knowledge and thinking. The first test, *Learning Mathematics for Teaching* (LMT), measured the development of content knowledge in elementary number and operations among pre-service teachers who took an IBL mathematics course targeted to this group. The other, the proof test, measured students' ability to evaluate and validate mathematical arguments, and here we could study differences between IBL and non-IBL students.

5.4.1 Summary and limitations of findings from the LMT test

The *Learning Mathematics for Teaching* (LMT) instrument was useful in detecting pre-service teachers' growth. These students clearly learned mathematical knowledge for teaching during their IBL course. Secondary pre-service teachers did better than elementary and middle school pre-service students on the pre- and post-test, but all three student groups scored better on average on the post-test.

Students with low pre-test scores improved more than did initially high-performing students. Other group differences were minor, indicating that pre-service teachers benefited from an IBL course regardless of their academic goals, teaching experience, or personal background.

Students' improvement in learning mathematics for teaching did not correlate particularly well with other measures of student learning outcomes. However, as noted above, initially weaker students improved more than those who started with better LMT test scores. This is similar to the finding reported in Chapter 3 (Section 3.1.6), weaker pre-service teachers, measured by their estimated GPA, tended to report slightly higher cognitive, affective and social gains on the SALG-M instrument than did other students. Together, these findings offer evidence that weaker students benefit more from an IBL course.

Because they are carefully developed and well-validated instruments, the LMT tests helped us to measure the kind of growth in mathematical knowledge needed for teaching K-12 mathematics. Our sample from two campuses was also large enough to detect real gains and differences among students. However, because IBL methods were used in all sections of the courses targeted to pre-service teachers that were available for this study, we had no opportunity to compare student learning with that in a traditionally taught course. This should be a significant element of any future study.

5.4.2 Summary and limitations of findings from the proof test

Many of the IBL courses focused on mathematical proving. The proof test was intended to measure students' ability to assess mathematical arguments on algebra, number theory and calculus. Overall, the results from the test indicate some strength in IBL students' ability to discriminate valid from invalid arguments, in their level of skepticism, and in their use of more expert-like criteria to judge an argument. However, the differences between IBL and non-IBL students were slight and most were not statistically significant. Our qualitative analysis of students' written reasoning reflected a greater reliance by IBL students on explanation and rigorous thinking. This contrasts with more frequent use of inadequate or false reasoning among non-IBL students.

Our ability to draw strong conclusions is limited by our sample of students. The students who volunteered to take the proof test were strong mathematics students, based on their self-reported grades and numbers of prior mathematics courses taken. We surmise that differences in the responses and reasoning of IBL vs. non-IBL students are less easily detected among this group than among lower-achieving or less experienced mathematics students—on this test or others we might devise. Students with high grades are by definition those who have previously succeeded under a wide range of instructional conditions; they are more likely to be teacher- and method-proof. Thus, while the test itself seems to be sensitive to differences in students' understanding, our sample is not optimized to detect group differences that might result from IBL instruction focused on proof processes. It is particular unfortunate that this sub-study includes few low-achieving students, since other types of evidence suggest that the impact of an IBL course is greater for lower-achieving and less mathematically experienced students (Ch. 3, Ch. 6, Ch. 8). Future data collection must necessarily focus on opportunities to administer this or other tests to *all* students—high- and lower-achieving—in a comparative setting with both IBL and non-IBL versions of the same course.

Lastly, this particular test is likely to be more sensitive in "introduction to proof" courses where enrollment is controlled or sequenced in such a way as to assure that most students have relatively little prior proof experience. In this sample, many students had proof experience already, and we cannot rule out that the test measured expertise developed in earlier courses.

5.5 References Cited

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

Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational Studies in Mathematics, 44*(1-2), 5-23.

Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics, 24*, 389–399.

Pfeiffer, K. (2009). The role of proof validation in students' mathematical learning. In Joubert, M. (Ed.), *Proceedings of the British Society for Research into Learning Mathematics, 29*(3) November 2009.

Selden, A., & Selden, J. (1999). *The role of logic in the validation of mathematical proofs.* Department of Mathematics, Technical report. Tennessee Technological University. TN: Cookeville.

Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education, 34*, 4-36.

Smith, J. C. (2006). A sense making approach to proof: Strategies of students in traditional and problem-based number theory courses. *Journal of Mathematical Behavior*, 25, 73-90.

Sowder, L., & Harel, G. (2003). Case studies of mathematics majors' proof understanding, production, and appreciation. *Canadian Journal of Science, Mathematics and Technology Education*, *3*(2), 251-267.

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 39(4)*, 431-459.

Weber, K. (2009). Proving is not convincing. *Conference on Research in Undergraduate Mathematics Education,* Raleigh, NC, February 26-March 1, 2009.

Chapter 6: Findings from Student Academic Records

6.1 Introduction

Grades are a traditional and standard way of measuring academic achievement. While instructor standards for assigning grades may differ, grades are nevertheless widely seen to hold a stable meaning across institutional contexts. Moreover, grades may reflect long-term changes in achievement, when improved learning habits and analytical thinking carry over to later courses. Students' choice of courses may also reflect sustained or lost interest in the discipline. While other parts of this study measure and compare immediate student outcomes of taking IBL and parallel non-IBL courses, institutional academic records data allow us to analyze student outcomes on a longer time scale as we look for patterns in student grades and class-taking choices. Thus, in this study we use grades as a longitudinal proxy for academic achievement. We use number of math courses taken subsequent to an IBL course and math major status as proxies of student interest and motivation in mathematics.

We obtained academic records from three campuses for all students enrolled in IBL and non-IBL sections of certain courses. We chose as our targets core IBL courses that had a well-established history, sizable enrollments, and available comparison groups. Defining our study samples as students enrolled in these courses (or their non-IBL counterparts), we then looked forward and backward in time to examine the impact of taking an IBL course on students' grades, course choices, and academic major choices.

With this academic records data we sought to address the following questions:

- What are the student outcomes from IBL instruction, as measured by grades, class-taking patterns, and academic major status?
- How do these student outcomes differ between IBL and non-IBL students?
- How do these student outcomes vary among student sub-groups?

The study methods are detailed in Appendix A6.[1] In short, we converted the raw data obtained directly from institutional records offices into standardized variables, counting enrollments in later math courses and computing grade-point averages (GPAs) for these courses. We constructed several different achievement measures based on grades. The *target course grade* offers a measure of immediate impact, but must be used with caution when comparing IBL vs. non-IBL courses, because, as instructors told us, the nature of student work and assessment systems tends to differ in IBL and non-IBL sections. The *average math grade in the semester immediately after the target course* (hereafter *next semester average grade*) offers a medium-timescale look at the student achievement. It has been used in the literature (Jensen, 2006) as a measure of teaching intervention making an impact on subsequent student achievement. Other

[1] http://www.colorado.edu/eer/research/documents/A6academicRecordsMethods.pdf

measures—such as average math grades across all semesters after the target course—provide a long-term picture of student achievement.

We also constructed several measures of motivation based on student class-taking patterns. We distinguished between the number of *required*, *elective*, and *IBL* courses taken *after* the target, because students may have different levels of interest in certain types of courses depending on their difficulty, preferences for teaching method, and importance for graduation. Detailed definitions of each variable are provided in Appendix A6.

Variables of average math grade and number of math courses taken *prior* to the target course serve as measures of students' *incoming* achievement and mathematical background, respectively. We use them to compare the IBL and non-IBL student groups *prior* to the target course: only if they are similar enough before the target course, can we confidently attribute the outcomes to the IBL intervention.

This analysis includes academic records from 3212 students from three courses at two campuses.

6.2 Comparing apples to apples: Design considerations for comparative records analyses

Because we are interested in the differences in outcomes for IBL and non-IBL students, insuring the comparability of both samples is paramount. First, to make an argument about an outcome of intervention, the two groups under investigation need to receive two different treatments. That is, IBL and non-IBL classes need to be different enough that it is reasonable to expect some student outcome differences. As the data in Chapter 2 show, that is generally the case.

Second, there must be an appropriate comparison group. On two campuses, non-IBL sections of some courses were taught simultaneously to the IBL sections, so the comparison group was students who enrolled in these sections and sections taught prior to establishment of the IBL Centers. Contemporaneous sections provide the best comparative case: students presumably had a choice of sections, but should be similar in background. On a third campus, there was no simultaneous section, so we attempted a historical comparison of IBL course sections with sections taught before the IBL Center was established. However, we learned that this course was in fact an ancestor to the IBL efforts at this campus, and was taught using essentially the same methods as the current course. Thus, the historical comparison was not appropriate, and we do not discuss student outcomes from this course.

Third, to make any argument that the intervention is responsible for the observed differences in outcome, both treatment groups need to be comparable *before* the intervention. That is, students enrolled in the IBL and non-IBL sections need to be similar prior to taking the target class, and if they are not, we need to control for those differences in our analysis. We took two approaches to this: in one case where incoming differences were drastic, we used sampling to create two matched groups. In more moderate cases, we used statistical techniques to control for incoming differences. For all the results discussed below, we used either one or both of these methods to control for incoming differences.

6.2.1 Constructing study samples

Since most of our measures relate to the courses taken and grades obtained *after* a particular course, we had to select situations where students were in fact likely to take some mathematics classes after that point. For example, we did not examine pre-service teacher outcomes, because we knew that these students rarely took math courses in excess of the required core, due to their very structured and rigorous curriculum. Moreover, we selected IBL target courses positioned early enough in the curriculum that students had the opportunity to take more classes after their IBL course. Thus, for example, we do not discuss the outcomes for one university's Real Analysis course because most students take it their senior year.

Following this logic, we focused on "target" course sections taught far enough back in time that students who took them had an opportunity to graduate, having taken all the subsequent courses they wanted or needed to take. This defined the study sample. Then we tracked the courses taken by those students and grades achieved from that point on, until graduation. As mentioned, students' grades in the target course itself might not be comparable to those in non-IBL course, due to differences in grading schemes and the nature of the work being graded. However, after the target course, both IBL and non-IBL students take the same courses and receive grades based on common standards. Therefore, it is certainly fair to compare the later outcomes of students with prior IBL or non-IBL backgrounds. Taking into account all these considerations, we were able to analyze student outcomes for three courses at two campuses, as listed in Table 6.1.

Table 6.1: Study Samples for Academic Records Analyses

Course name	Course code	Method of control for incoming difference	Total sample size	IBL sample size	Non-IBL sample size
University L Mid-level course	L1	Statistical control	1341	211	1130
University L Upper-level course	L2	Statistical control	909	123	786
University G Introductory course	G1	Sampling, statistical control	962; 197 after sampling	49	98

The sample size for analyses of individual variables is generally smaller than the values in Table 6.1, because not all students complete courses of all types subsequent to their target course. Non-IBL samples are generally larger because of the limited offerings of IBL sections.

6.2.2 Prior to the target course, IBL students had generally earned higher grades and taken fewer math classes than their non-IBL peers.

We use two incoming measures to compare students who enrolled in IBL and non-IBL sections of a course: the number of math courses taken and average math grades prior to the target course. Students enrolled in IBL sections of L1 had up to that point taken fewer math courses and obtained higher average grades than their non-IBL peers. Those differences in means are statistically significant. We dealt with these incoming differences by employing statistical procedures that control for them in analyzing outcomes. Students enrolled in the IBL section of L2 were not statistically significantly different from their non-IBL peers in the two incoming variables. (Of course, there may be other differences that we cannot measure.)

Students in the IBL and non-IBL sections of G1 were initially drastically different on the two incoming measures and other suitable indicators of background for this entry-level course, including high school GPA and college admission test scores. Since IBL sections of G1 are honors courses, students are invited into them based on their stellar record of prior achievement. The non-IBL sections, on the other hand, admit students of all levels of ability and mathematical achievement. Thus, in order to accomplish a valid comparison, we sampled students from the larger non-IBL group who closely match the select IBL population on multiple criteria, including college admission test scores, high school GPA, major, academic year, gender, and race. After sampling, the difference in average prior grades disappeared. However, the difference in the number of math classes taken prior remained statistically significant, so we applied statistical procedures to control for this difference in analyzing outcomes.

It is interesting that the difference in the number of classes taken prior to target remains significant after sampling, which includes matching students on the academic year criterion. That is, IBL freshmen were matched with similar non-IBL freshmen, but the former still took fewer math courses prior to G1 than did the latter. This is important evidence of selection effects – students testing directly into G1 are hand-picked into the IBL sections, while other high-achievers who still need some prerequisites end up in the non-IBL sections.

6.3 Comparison of IBL and non-IBL students: Differences in grades and number of subsequent courses

We use subsequent average grades as proxies of achievement, and numbers of subsequent math courses as proxies of student motivation after taking the IBL or non-IBL section of a course. Below we discuss some key differences in these achievement and motivation outcomes between students with IBL and non-IBL backgrounds. The statistical significance of differences in group means is indicated in the tables. In later sections, we examine sub-group differences.

6.3.1 Students achieved equal or higher average math grades after taking an IBL class.

Results for the grade variables for all three courses (L1, L2, G1) are shown in Table 6.2. In course L1, students with IBL backgrounds performed better on all grade measures than their non-IBL peers. However, in only one case is this difference statistically significant: IBL students scored significantly higher in subsequent IBL courses—almost half a grade higher than their non-IBL peers.

In course L2, students with IBL background also scored higher or the same on all grade measures compared to their non-IBL peers. Although the differences for L2 generally follow the same trend as for L1, none of them are statistically significant, perhaps because of the smaller sample sizes and L2's position later in University L's curriculum.

Table 6.2: Means for Average Grades for IBL and Non-IBL Students

Variables based on average grades in subsequent math classes	**IBL mean**	**non-IBL mean**	**IBL N**	**Non-IBL N**	**Sig.**
L1					
Target course grade	2.66	2.59	204	1077	
Next semester average grade	2.96	2.79	89	526	
Average grade in subsequent required courses	2.80	2.63	104	499	
Average grade in subsequent elective courses	2.95	2.83	130	725	
Average grade in subsequent IBL courses	2.98	2.53	30	80	*
L2					
Target course grade	2.52	2.48	117	747	
Next semester average grade	2.80	2.59	41	326	
Average grade in subsequent required courses	2.56	2.27	22	130	
Average grade in subsequent elective courses	2.69	2.67	71	446	
Average grade in subsequent IBL courses	2.56	2.57	7	23	
G1					
Target course grade	3.34	3.18	47	98	
Next semester average grade	3.20	3.00	39	68	*
Average grade in subsequent required courses	3.14	3.05	40	70	
Average grade in subsequent elective courses	2.98	3.03	16	20	
Average grade in subsequent IBL courses	3.47	3.41	35	3	

p < .05, p < .01**, p < .001****

Finally, in course G1, students with IBL background likewise scored higher or about the same as their non-IBL peers on all measures. Although the differences for G1 mostly point in the same direction as for L1 and L2, only one is statistically significant. This is likely due to the closely matched samples: both the IBL sample and the non-IBL student sub-sample include very high-achieving self-motivated students who receive high grades regardless of the teaching method. However, it is important to point out that IBL teaching methods do not harm these high-achieving students. The finding of (almost) no difference among top students is corroborated by other parts of the academic records study, as we discuss below. The test data in Chapter 5 and some survey data in Chapter 3 point to this as well: there is less difference in achievement outcomes for the strongest IBL and non-IBL students, and more for lower-achieving students.

Despite the close matching, one difference is statistically significant: IBL students received a higher grade the semester immediately after G1 than did their non-IBL peers. Since top students' achievement is generally less sensitive to the teaching method, this seems to be a strong finding. However, for IBL students, the next semester grade is likely to include a grade in another IBL course, as G1 is a first course of a three-term sequence available in IBL. Since IBL and non-IBL courses are often graded differently, this finding is not as definitive as it may appear.

6.3.2 IBL's effect on students' motivation to pursue further math courses was not consistent.

While the achievement outcomes consistently point in the direction of IBL students benefitting more than their non-IBL counterparts, the picture is not as clear with the measures of motivation, as reflected in course-taking choices (Table 6.3).

Table 6.3: Mean Number of Subsequent Math Classes for IBL and Non-IBL Students

Number of new subsequent math classes taken	**IBL students mean**	**non-IBL students mean**	**IBL sample size**	**Non-IBL sample size**	**Sig.**
L1					
All math classes	2.25	2.31	204	1077	
Required classes	0.60	0.51	204	1077	*
Elective classes	1.64	1.80	204	1077	
IBL classes	0.14	0.06	204	1077	***
L2					
All math classes	1.36	1.47	117	747	
Required classes	0.02	0.04	117	747	
Elective classes	1.34	1.43	117	747	
IBL classes	0.02	0.02	117	747	
G1					
All math classes	3.29	2.62	47	98	
Required classes	2.09	1.96	47	98	
Elective classes	1.20	0.66	47	98	
IBL classes	1.21	0.03	47	98	***

$p < .05$, $p < .01$**, $p < .001$****

For course L1, overall students tended to take about as many new math courses after taking the IBL section as their non-IBL peers. They took statistically significantly more new required courses and new IBL courses. Students with IBL background took more than twice the number of IBL classes taken by their non-IBL peers. On the other hand, student with IBL background seem to have taken slightly fewer elective courses than their non-IBL counterparts. While this result is not statistically significant, it is perhaps unsurprising. As student interviews confirm (Ch. 7), IBL courses involve a lot of hard work for everyone, but especially for weaker students. Thus, the effort required in an IBL course may deter students from taking more elective courses, which are not essential to their academic advancement in the mathematics major. As discussed above, IBL students took statistically significantly more courses that were essential to their completion of the major (required) and also chose more courses using inquiry than did their non-IBL peers.

For course L2, there is no statistically significant difference in the course-taking patterns of students who took IBL and non-IBL sections. In most cases, IBL students took slightly fewer new math classes than their non-IBL counterparts. However, all the means for both groups are quite low, since this class is generally taken late in students' academic career. That is, neither IBL nor non-IBL students took many math classes afterwards.

As Table 6.2 (Section G1) shows, G1 students display the most consistent trend in motivation. G1 IBL students took more subsequent math courses on all measures than their non-IBL peers, although only one of these differences is statistically significant. IBL students were especially likely to take subsequent IBL courses compared with their non-IBL peers. G1 students had the opportunity to take more IBL courses, since G1 is the first course in a three-term sequence available in IBL. Since at least two more IBL classes were available, the mean of 1.21 means some student s did not finish the sequence. These may be majors whose degree requirements stop at G1. IBL students' consistent motivation to take more math classes was perhaps related to the high-achieving status of these students. Top achievers may be less discouraged by the hard work involved in IBL course, retaining and even increasing their motivation to take further math classes. It is significant, however, that a traditional approach to the course did not generate similar interest in pursuing mathematics among high-achieving mathematics students.

We also found no significant difference between IBL and non-IBL students on other measures of motivation and interest: adding or dropping the mathematics major. IBL students did not add math major in higher numbers or drop it in lower numbers than their non-IBL peers.

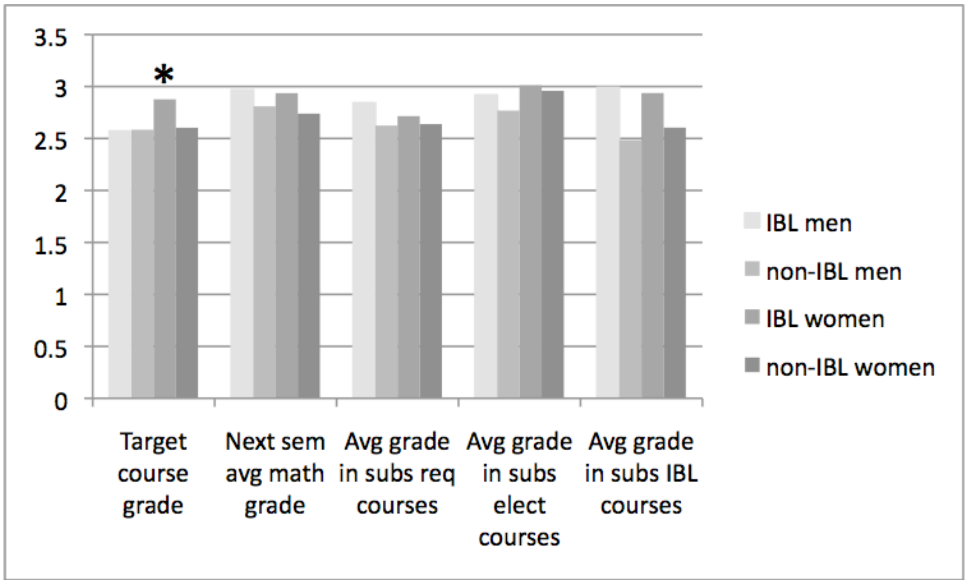
In sum, we observe no clear trend for the impact of IBL courses on students' motivation across the three courses analyzed. Overall, it seems that IBL students may be somewhat more motivated to take further math courses, especially those required for completing a math degree and those taught with IBL methods. This is especially true for the high-achieving students in course G1. We further discuss the effect of achievement status on motivation to take more courses below.

6.4 Differences by gender: Subsequent grades

Since gender is often cited as a factor in mathematical achievement outcomes, and because benefits of IBL for women have been observed in other sub-studies (Ch. 3), we compared subsequent grades for men and women. In this section, we discuss some of the key achievement differences between IBL and non-IBL men and women. Results for courses L1, L2 and G1 are shown in Figures 6.1, 6.2 and 6.3.

The sample sizes in these figures vary significantly from variable to variable, as not all students enrolled in all kinds of courses. For specific sample sizes, see Appendix A6.

Figure 6.1: Means in Average Subsequent Grades by Student Gender & IBL Status, for Course L1



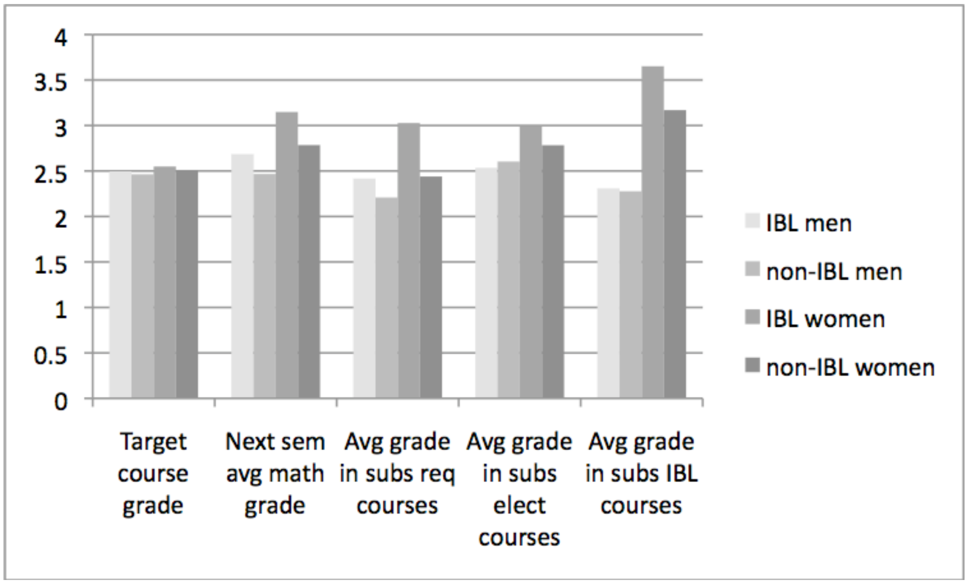
As Figures 6.1-6.3 show, IBL instruction was beneficial to both men and women. While most differences in means are not statistically significant, it is important that both IBL men and IBL women achieved the same or higher average subsequent grades than did their non-IBL counterparts in most cases. This trend holds across the three target courses discussed, with one exception: G1 women with IBL background received lower average grades in subsequent elective courses than did non-IBL women (Figure 6.3).

6.4.1 IBL women earned higher or equal average grades than non-IBL women in later math courses.

As Figures 6.1-6.3 show, women with IBL background mostly received higher average grades in subsequent courses than their non-IBL peers. There are two exceptions to this trend: G1 IBL women scored about the same in subsequent required courses, and lower than non-IBL women in subsequent elective courses (Figure 6.3). But these differences are not statistically significant. Only one difference is statistically significant here: L1 women with IBL background scored significantly higher in the target course than non-IBL women.

Despite the overall lack of statistical significance, we see a rather consistent trend (especially in L1 and L2) of IBL women scoring higher than their non-IBL counterparts. This trend is important because we use two different kinds of achievement measures. The target course grade is a grade received in a IBL or non-IBL section of the target course. Since assessment methods often vary with teaching practices, grades in IBL classes may measure and emphasize different things than grades in non-IBL courses. Indeed, in interviews instructors told us that they graded IBL courses differently—for example, emphasizing class participation more than in their lecture-based courses. Thus, the fact that IBL women achieved better grades in the target course may not point to real differences in performance. However, measures that use only subsequent grades compare grades received in the same kinds of courses and based on the same kind of assessment. The fact that IBL women mostly got higher grades on those measures points to lasting improvement in their mathematical performance.

Figure 6.2: Means in Average Subsequent Grades by Student Gender & IBL Status, for Course L2



6.4.2 IBL men earned equal or higher average grades than non-IBL men.

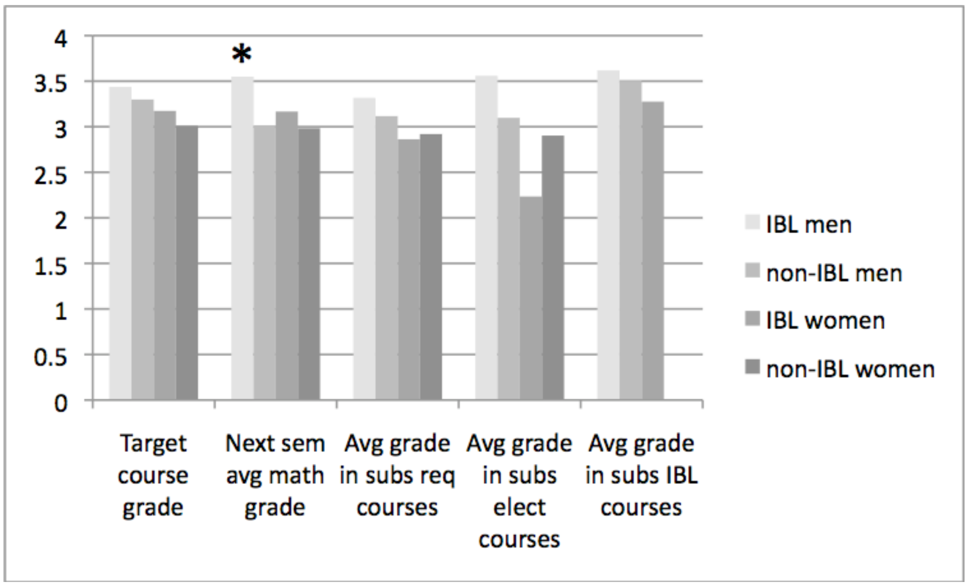
As Figures 6.1-6.3 show, IBL men earned the same or higher average grades than did their non-IBL counterparts on all measures. Most of these differences are not statistically significant, but one difference is: G1 men with IBL background received a significantly higher average math grade the semester after taking the target course, over half a grade higher than their non-IBL counterparts. Overall, the data for IBL men show a rather consistent trend of slightly higher grades on most measures compared with their non-IBL peers.

6.4.3 IBL instruction closed some of the gender gap present in traditional math instruction.

As Figure 6.1 shows, after taking L1, non-IBL women lagged behind their male classmates on some measures of achievement. They received about the same grades as men in the target course itself but earned lower average math grades the semester immediately after taking L1. On the other hand, after taking L1, IBL women did as well or better on these indicators than their male classmates. They received higher grades then IBL men in the target course itself and about the same grades the next semester. Based on these indicators, IBL instruction appeared to enable women, who are often disadvantaged by traditional mathematics instruction (Linn & Kessel, 1995), to catch up to or even surpass men in achievement.

Other measures of achievement yield less clear results. While non-IBL L1 women got about the same average required grades and slightly higher average IBL grades compared to their male class mates, IBL women did not fully catch up to their male classmates on those measures. On the other hand, IBL women received higher average grades on those measures than did non-IBL women. Thus, IBL women earned good grades in absolute terms, but they still somewhat trailed behind IBL men, who got even better grades. Examining average grades in elective courses, both IBL and non-IBL women surpassed their male classmates in achievement.

Figure 6.3: Means in Average Subsequent Grades by Student Gender & IBL status, for Course G1



As Figure 6.2 shows, for course L2, there are no statistically significant differences in the average grades between IBL and non-IBL men and women. Interestingly enough, both IBL and non-IBL women got as good or higher average grades as their male classmates in and after this course. This is due to the fact that this course is situated rather far along in University L's curriculum, so most students take it during their senior year. By this point in their academic career, women who had a hard time with traditional mathematics instruction may have left the major to pursue other interests (Seymour & Hewitt, 1997). The women who still remain in the major at this point are survivors, stronger or more persistent students who are "instruction-proof." Thus, regardless of the use of IBL or non-IBL methods, women in and after L2 did as well or better than their male classmates on achievement indicators.

On the other hand, as Figure 6.3 shows, both IBL and non-IBL women received lower average grades in and after course G1 than their male classmates. This could be because G1 is situated very early on in University G's mathematics curriculum; for many students, it is the first or second math class they take. Women may be less sensitive to the teaching method than to the general challenges of college-level math, especially if women enter college already underestimating their mathematical ability or sensitive to gender stereotypes about mathematics (Pronin, Steele, & Ross, 2004; Seymour, 1995).

6.5 Differences by prior achievement: Grades

In this section, we discuss some of the key achievement differences between IBL and non-IBL students within three groups divided by their mathematics GPA previous to the IBL course: low, medium, and high.[2] We decided to explore the relationship between prior achievement and student outcomes after noting the strong correlations between those variables (for details see Appendix A6). For the sake of simplicity, however, we present the results broken out by three prior achievement levels.

For course L1, the subgroups are defined by the following mathematics GPA divisions: low, GPA of 0-2.5; medium, GPA of 2.5-3.4; and high, GPA of 3.4-4.0. For course L2, the divisions are: low, GPA of 0-2.4; medium, GPA of 2.4-3.4; and high, GPA of 3.2-4.0. These cutoffs are slightly different because they were developed empirically, with the goal of dividing the students into three equal-sized groups. For course G1, we do not conduct this analysis, because G1 students were all relatively high-achieving, given our sampling procedures. Moreover, most G1 students had no prior mathematics GPA as G1 was their first college mathematics course.

6.5.1 IBL students with previously low GPA mostly earned higher average grades in subsequent courses than did low-achieving non-IBL students.

As Figure 6.4 shows, IBL students who carried a low GPA prior to taking L1 scored higher on all achievement measures than did their non-IBL counterparts. IBL low-achievers received statistically significantly higher average grades in subsequent required and IBL courses as compared with non-IBL low achievers. While differences on other measures are not statistically significant, they all point in the same direction, with grade benefits to IBL students with low prior GPA.

On the other hand, this trend is less clear in the achievement outcomes of L2 students. As Figure 6.5 shows, IBL students who carried a low GPA prior to taking L2 received slightly higher grades the semester immediately following L2 and higher average grade in subsequent required courses, versus their non-IBL counterparts. However, low-achieving IBL students received slightly lower grades than low-achieving non-IBL students on all other metrics of achievement. None of these differences are statistically significant, and most sample sizes are small. Thus, the inconsistency in L2 results does not invalidate the trend established by the strong, statistically significant results for L1: low-achieving IBL students scored higher than their non-IBL peers, especially on subsequent required and IBL courses.

---

[2] These groups differ from the achievement groups used in Chapter 3. Here we use only mathematics grades from academic records data, instead of self-reported general college GPA.

Figure 6.4: Means for Average Subsequent Grades by Student Prior GPA & IBL Status, for Course L1

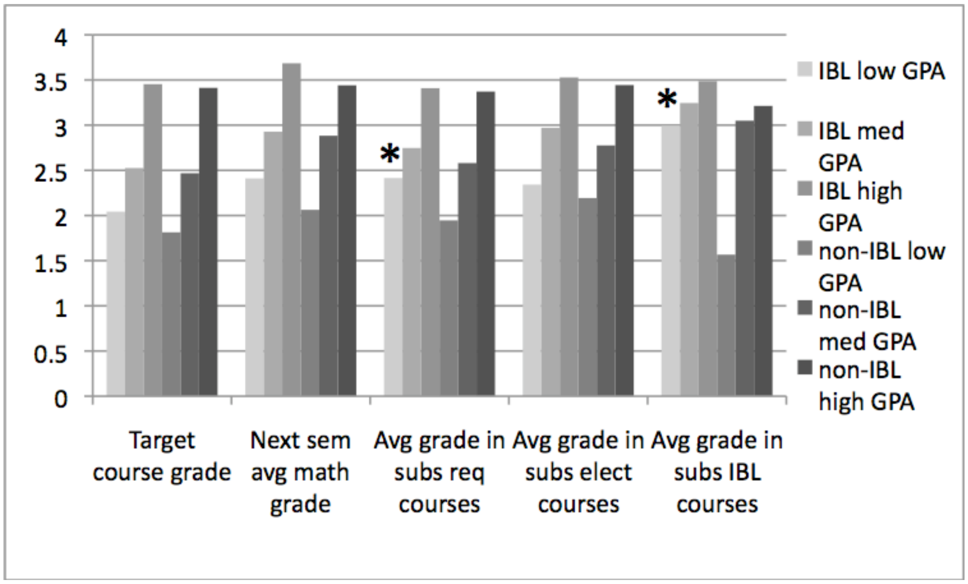
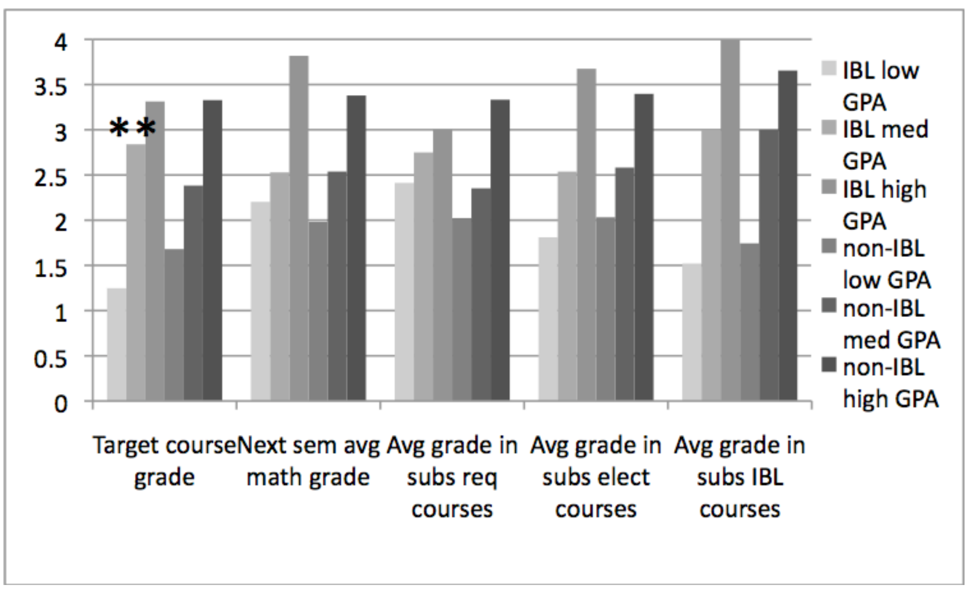


Figure 6.5: Means for Average Subsequent Grades by Student Prior GPA & IBL Status, for Course L2



6.5.2 IBL students with previously medium GPA earned equal or higher average grades than medium-achieving non-IBL students.

As Figures 6.4 and 6.5 show, IBL students who carried a medium GPA prior to taking either L1 or L2 scored about the same or higher on all achievement measures compared with their non-IBL peers. This trend holds across both classes. One of these differences is statistically significant: L2

IBL medium-achievers received significantly better target course grades, almost half a grade higher than non-IBL medium-achievers.

Once again, it is important that the trend benefitting IBL medium-achievers is consistent across two types of outcome variables. While the target course grade may not be directly comparable between IBL and non-IBL students, subsequent grade measures compare the grades of the same kinds of courses. These results points to consistent, lasting changes in mathematics achievement for medium-GPA IBL students.

6.5.3 IBL students with previously high GPA earned equal or higher average grades than high-achieving non-IBL students.

As Figures 6.4 and 6.5 show, IBL students who carried a high GPA prior to taking either L1 or L2 scored about the same or higher than their non-IBL peers on all achievement measures except one. In course L2, high-achievers with IBL background received lower average grades on subsequent required courses than did non-IBL high-achievers. None of the differences, however, are statistically significant. The lack of statistical significance is perhaps due to the high-achieving status of these students. These students may be somewhat "instruction-proof": they receive good grades regardless of the teaching method. Once again, these findings are consistent with test results in Chapter 5 and some of the survey results in Chapter 3, which suggest that there is less difference in achievement outcomes for the strongest IBL and non-IBL students. Importantly, IBL methods do not harm these top students.

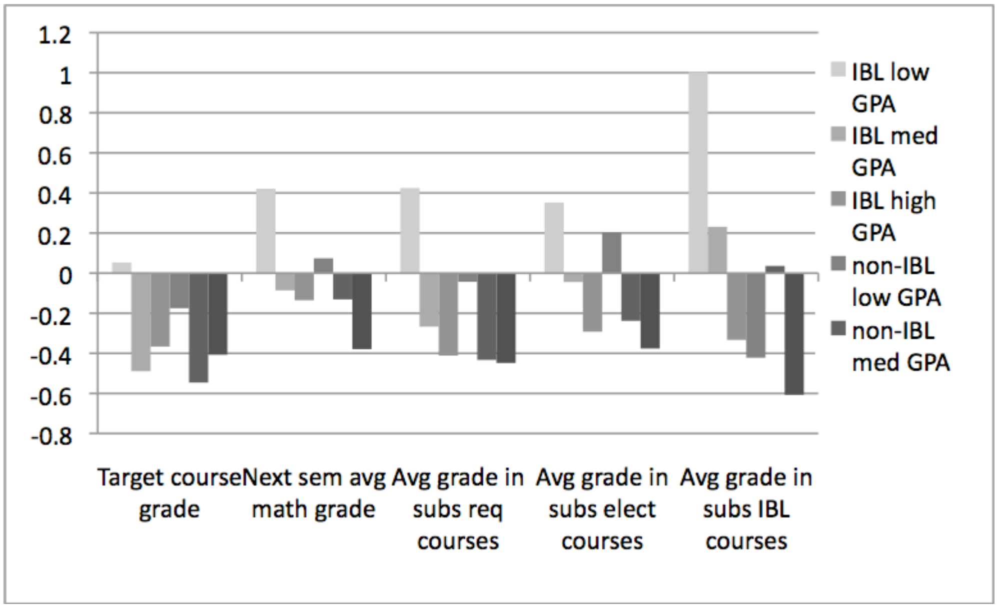
6.5.4 IBL instruction enabled low-achieving students to improve their grades, as a group.

As Figures 6.4 and 6.5 show, students with prior medium GPAs scored higher on subsequent courses than did students with prior low GPAs. The students with prior high GPAs achieved higher grades than medium-achieving students. Thus, the students who did well before the target course continued to do well, and the students who struggled before taking L1 or L2 continued to struggle in subsequent courses. This trend holds for both IBL and non-IBL sections of L1 and L2. However, there is a marked difference in the subsequent grades of low-achieving students with IBL and non-IBL background in L1, relative to their prior achievement as a group.

Figure 6.6 displays the change in average student grades from before to after L1, broken out by achievement group. For example, the prior GPA average for the low-achieving group is 1.99. To assess how average grades of low-achievers changed from before to after L1, we subtracted 1.99 from the low-achievers' average grades after L1 (i.e., from the means shown in Figure 6.4). We followed the same procedure for medium- and high-achieving groups, subtracting their prior group average grades—3.01 and 3.82, respectively—from their group average grades after L1.

As the resulting Figure 6.6 shows, the changes in average grades of IBL low-achievers from before to after L1 are consistently positive. Thus, IBL low-achieving students got consistently higher average grades in subsequent courses than their pre-L1 group average grades. On the other hand, non-IBL low-achievers earned (modestly) higher average grades than their prior GPA only on two measures. That is, IBL instruction enabled low-achieving IBL students to earn consistently *higher average grades than would be expected by their past performance*, closing some of the achievement gap that separated them from medium and high-achievers in traditional courses.

Figure 6.6: Change in Grades (Before/After L1), by Student Prior GPA and IBL Status



Hence, so far we have seen that the IBL instruction strongly benefited the achievement of women and low-achieving students, without harming men and high achievers.

6.6 Differences by gender: Students' motivation to pursue further math courses

In this section, we discuss the key differences in numbers of subsequent courses taken between IBL and non-IBL men and women. Again, we use number of subsequent courses taken as proxy for student motivation in mathematics.

6.6.1 IBL women took as many or more subsequent required and IBL courses than did non-IBL women.

As Figures 6.7-6.9 show, IBL women took as many or more subsequent required mathematics courses than their non-IBL counterparts. Women with IBL background in L1 took statistically significantly more required courses after L1 than non-IBL women (Figure 6.7). Women with IBL background in courses L2 and G1 on average took as many subsequent required courses as their non-IBL counterparts. These results are not statistically significant.

Figures 6.7-6.9 also show that IBL women took as many or more subsequent inquiry-based courses than their non-IBL counterparts. Women with IBL background in L2 on average took as many IBL courses after taking L2 as non-IBL women. Women with IBL background in L1 and G1 took more subsequent IBL courses than their non-IBL peers. Only for G1 this difference is statistically significant.

Required courses are the core courses that must be completed to graduate with a mathematics degree. While students can sometimes select the required courses from a list of predefined choices, the number of such choices is fairly small in each mathematics program. Thus, it is not clear what taking more required courses tells us about IBL women's motivation.

Figure 6.7: Mean Number of Subsequent Courses by Student Gender & IBL Status, for Course L1

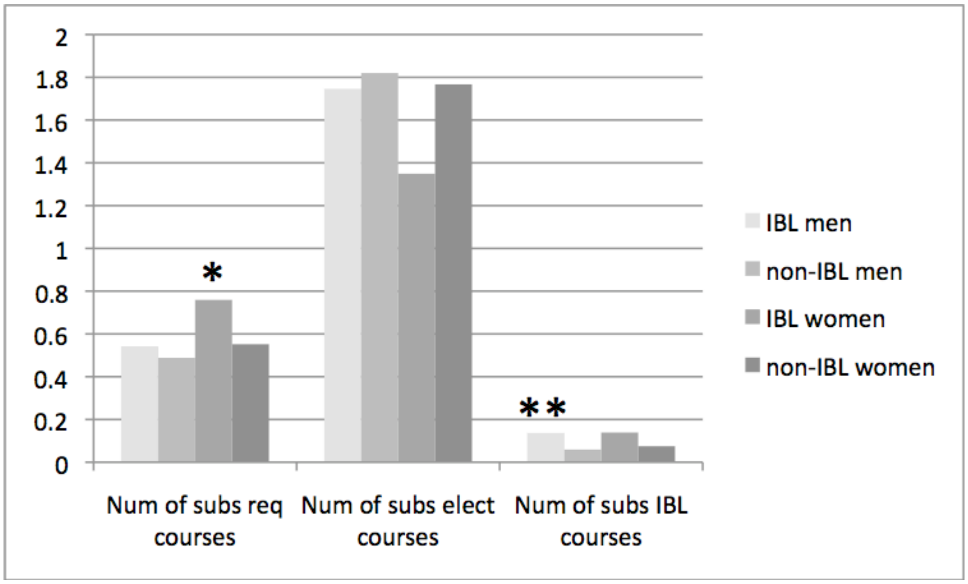
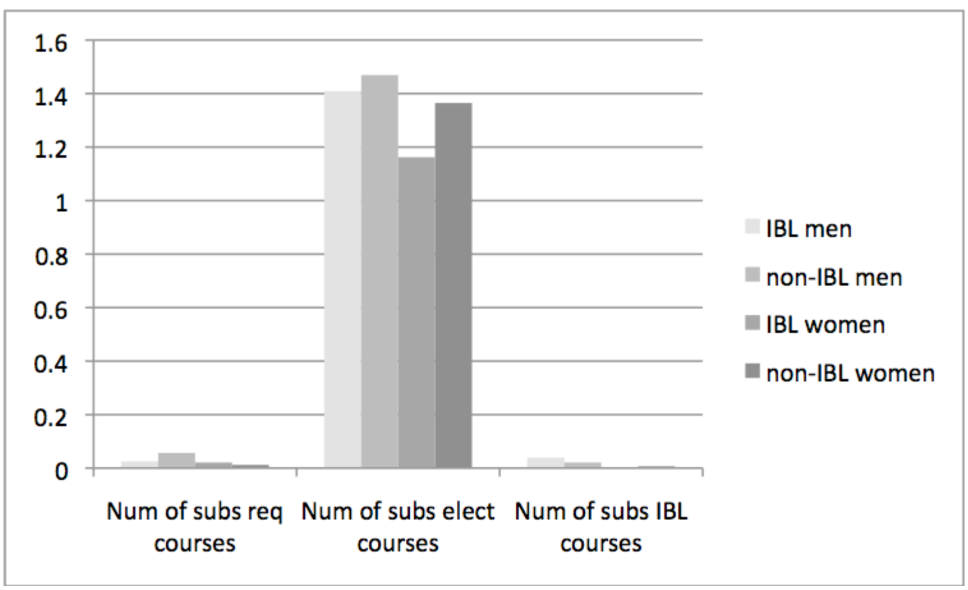


Figure 6.8: Mean Number of Subsequent Courses by Student Gender & IBL Status, for Course L2



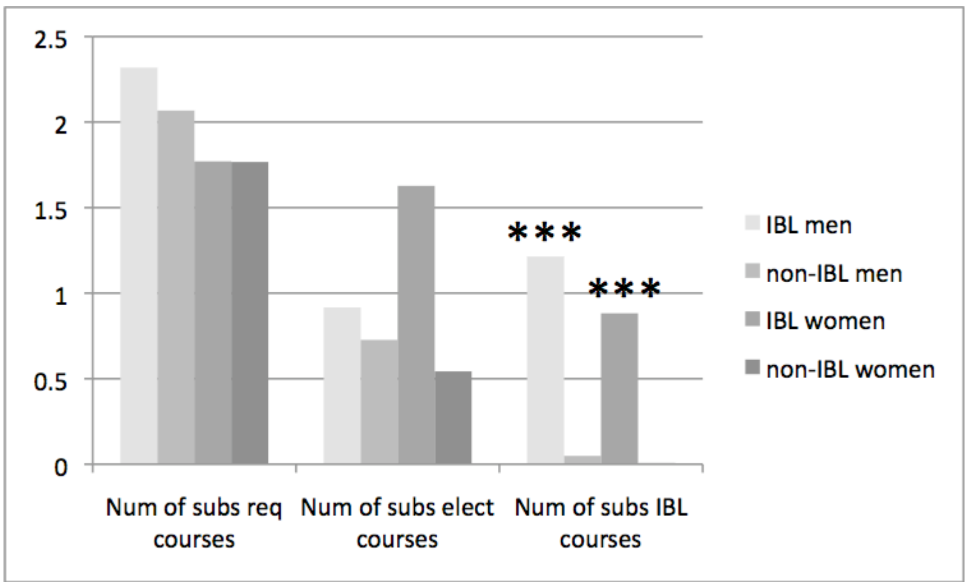
On the other hand, women with IBL background in L1 and L2 took fewer subsequent elective courses than non-IBL women. We discuss this further in Section *6.5.3* below, addressing both men and women.

6.6.2 IBL men took more subsequent IBL courses than did non-IBL men.

As Figures 6.7-6.9 show, IBL men took more subsequent IBL courses than did their non-IBL counterparts after all three target courses. While this difference is not statistically significant for L2, men with IBL background in L1 and G1 took statistically significantly more subsequent IBL courses than non-IBL men. This difference is especially drastic for G1, where at least two more IBL courses where available after the target course.

Men with IBL background in courses L1 and G1 also took more subsequent required courses than their non-IBL counterparts (L2 IBL men took slightly fewer). These differences are not statistically significant, but they point in the same direction as the results for subsequent IBL courses.

Figure 6.9: Mean Number of Subsequent Courses by Student Gender & IBL Status, for Course G1



Overall, the use of IBL teaching methods seems to have positively influenced men's motivation to take more required and IBL courses. This is especially the case for subsequent IBL courses, suggesting that many students found one IBL experience sufficiently appealing to pursue another. IBL men got the biggest and the most significant boost in motivation to take further math courses.

6.6.3 G1 IBL students took more electives than non-IBL students, while L1 and L2 IBL students took fewer.

For both L1 and L2 courses, both men and women with IBL background on average took fewer elective courses than their non-IBL counterparts (Figures 6.7 and 6.8), although these differences are not statistically significant. IBL courses involve a lot of hard work for everyone, but may be especially challenging for weaker students. Thus, the effort involved in an IBL course might not particularly motivate students to take more elective courses, unessential for graduation with a math degree. As Figures 6.7 and 6.8 show, this result is especially apparent for women.

However, for the G1 course, results for both men and women with IBL background display an opposite trend. While this difference is not statistically significant, G1 IBL students took more subsequent elective courses then their non-IBL peers. This could be due to the timing of the intervention: G1 is the first course in a three-term sequence that students generally take in their first or second year in college. This is an early intervention, compared to L1 and L2, which are generally taken by juniors and seniors. Some of the survey data in Chapter 3 support the notion that earlier exposure to IBL methods makes a higher impact on students.

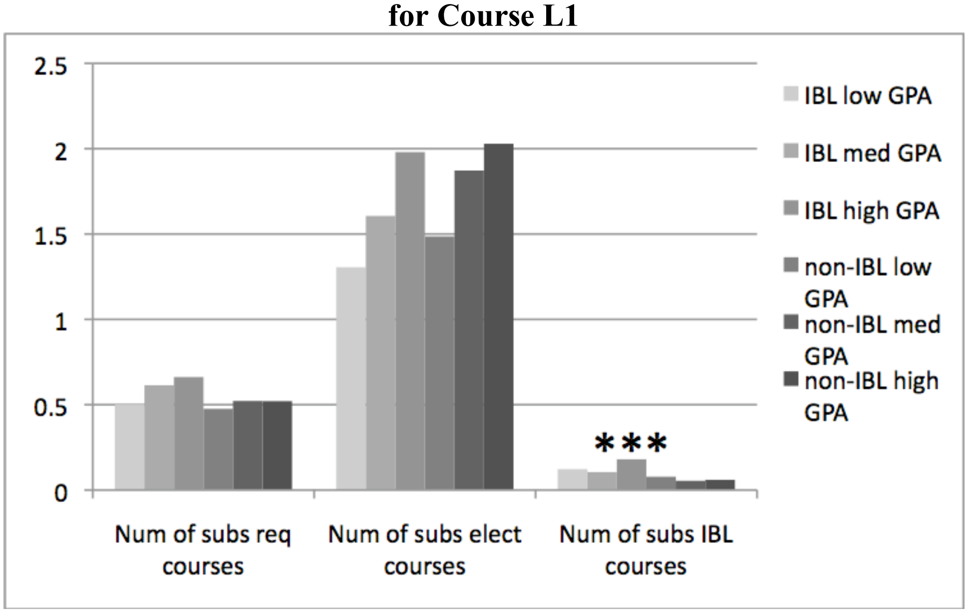
Hence, it is possible that G1 students, who were exposed to IBL early on, were encouraged by intellectual stimulation of the method, as opposed to being discouraged by the hard work involved. As mentioned in Section *6.3.2*, as high achievers, G1 students may also be resistant to the elective discouragement effect seen in L1 and L2.

6.7 Differences by prior achievement: Students' motivation to pursue further math courses

In this section, we discuss the key differences in numbers of subsequent courses taken between IBL and non-IBL students who were previously low-, medium-, or high-achieving. Students in our G1 data set are all rather high-achieving, so this analysis was not appropriate. Outcomes for L2 students are not statistically significant and do not show any clear trends. Thus, in this section we discuss motivation outcomes, in terms of course-taking patterns, of L1 students only.

6.7.1 IBL's effect on low- and medium-achieving students' motivation to pursue further math courses is not consistent.

Figure 6.10: Mean Number of Subsequent Courses by Student Prior GPA and IBL Status, for Course L1



As Figure 6.10 shows, L1 IBL students with previously low and medium GPAs took more subsequent required and IBL courses than their non-IBL peers. However, these differences are not statistically significant.

Moreover, the trend of IBL benefitting low- and medium- achieving students is absent for the variable measuring number of subsequent elective courses. Students with low and medium prior GPAs on average took fewer elective courses than their respective non-IBL counterparts. Once again, this could be due to the amount of work involved in participation in an IBL class, which may be discouraging to weaker students.

Hence, we see no clear trend in the motivation outcomes for low- and medium-achieving students, despite the fact that the achievement outcomes for these IBL students were more positive. IBL students with previously low and medium GPA took more of the classes needed for graduating with math degree and more courses taught with IBL methods, but fewer elective courses that are optional. And none of these differences are statistically significant.

6.7.2 IBL students with previously high GPAs took more subsequent IBL courses than non-IBL high-achieving students.

As Figure 6.10 shows, L1 students with high prior GPAs took more subsequent IBL courses than do their non-IBL counterparts. This difference is highly statistically significant. L1 high-achievers with IBL background also took more subsequent required courses, and about as many elective courses as the non-IBL high achievers. These differences are not statistically significant, but they point in the same direction as the results for subsequent IBL courses.

As discussed in Section *6.3.2*, G1 students, who are all fairly high-achieving, also took notably more subsequent IBL courses than their non-IBL peers. This difference is highly statistically significant.

These results suggest that IBL methods provide high-achieving students with a strong boost in motivation to take further math courses, especially courses taught using IBL teaching practices.

6.8 Conclusions and Limitations of the Academic Records Study

Overall, there are several main trends apparent from our analysis. IBL teaching methods seem to have benefited all the students in some manner, with different subgroups reaping different benefits. In terms of academic achievement, some groups of students who are usually under-served by traditional mathematics instruction seem to have benefited most. IBL women earned better grades than their non-IBL counterparts, closing some of the achievement gap that separated them from their male classmates in traditional courses. Likewise, students with weak and modest records of prior academic achievement benefited in lasting ways from IBL instruction, showing improvements in later course grades. The consistency of these differences across multiple measures is striking, even though not all results are statistically significant.

Equally important is that the achievement of men and of high-GPA students was not harmed by the IBL methods. Moreover, these students benefited in a different way. Men and students with high prior GPA apparently got a boost in motivation to take further math courses, especially those taught using IBL teaching methods.

The results are rather sensitive to the target course chosen and its placement in the curriculum. The differences in outcomes for courses L1 and L2 demonstrate that quite well. Both course populations are very similar—in fact, many students who have taken L1 continue to take L2. However, all the results discussed above are much stronger for the L1 population than for L2. This has to do with students' needs and opportunities to take further mathematics courses, which depends on the placement of the course in the curriculum. Students taking L2 are well along in their academic career and have little need or opportunity to take more mathematics courses—thus the outcome measures available from academic records are less sensitive to the possible influence of IBL experiences on these students.

Thus, as tempting as it may be, comparison of the outcomes of the three courses would not be fair. Courses L1, L2, and G1 occur at different points in the curriculum and hence offer students different opportunities to take further math courses. Differing degree requirements also affect student choices. For example, science students taking G1 may stop taking math once they meet their science degree requirements, whether or not they enjoy and value their mathematics coursework. Unfortunately, the sample sizes available are too small to support analysis of this issue.

Another issue that invalidates direct comparison among the outcomes of L1, L2, and G1 is student selection. Both self-selection (student choice) and institutional selection (advising and/or prerequisite requirements) play out differently for each course. For courses L1 and L2, there is no institutional selection into the IBL sections, and self-selection is present but not very strong. Often students do not know if they are signing up for the IBL or non-IBL sections of these courses. On the other hand, in G1 we observe strong instructional selection: first-year students testing directly into G1 based on AP scores and high school background are invited into IBL sections. This selects for young and high-achieving students. Thus, the outcomes for the IBL G1 course are quite different from the other courses: we observe high average grades and high pursuit of further math courses. Moreover, boosts in motivation that G1 students got from their IBL experience are the most consistent out of the three data sets. This is consistent with some findings from Chapter 3, which suggest that IBL interventions have the most impact early in students' academic career. However, due to the elite nature of this introductory course, the impact of early intervention cannot be disentangled from that of the highly select population.

For many outcome variables, the differences in IBL and non-IBL means appear to be large, but are not statistically significant. This is due to the high level of variance in the data. In other words, the range of outcomes is rather wide for this student population. Such variability of the data makes it harder to achieve statistical significance, but also means that the observed outcomes that were found statistically significant are more robust.

6.9 References Cited

Jensen, J. (2006). Surprising effects of inquiry based learning. *Texas College Mathematics Journal, 3*(1), 1-9.

Linn, M., & Kessel, C. (1996). Success in mathematics: Increasing talent and gender diversity among college majors. In Dubinsky, E., Schoenfeld, A.H., Kaput, J., eds., *Research in Collegiate Mathematics Education II*, (pp. 101-144). Providence, RI: American Mathematical Society.

Pronin, E., Steele, C. M., & Ross, L. (2004). Identity bifurcation in response to stereotype threat: Women and mathematics. *Journal of Experimental Social Psychology 40*(2), 152-168.

Seymour, E. (1995). The loss of women from science, mathematics, and engineering undergraduate majors: and explanatory account. *Science Education, 79*(4), 437-473.

Seymour, E., & Hewitt, N. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview.

Chapter 7: Findings from Student Interviews

7.1 Introduction

Interviews with students allow us to directly investigate students' experiences in IBL mathematics courses and their perspectives on classroom learning and teaching processes. The student interviews address the following questions:

- What do students gain (or not) as a result of IBL instruction?
- How do students describe the teaching and learning practices in their classrooms? And how do these practices affect their learning?
- What difficulties do students encounter when learning with IBL methods? What resources are needed to overcome these difficulties?
- Do some students benefit more than others from IBL methods?

Compared to pre-structured surveys or tests, interviews allow students to offer their own views and ideas, spontaneously raise new issues, and explain their observations. Thus these data serve to illuminate new issues as well as to confirm and clarify results from the other sub-studies. This chapter focuses on the learning gains that students described as a result of their IBL mathematics class, and on their views of the processes that led to these gains.

Briefly, we interviewed 68 students who had taken an IBL mathematics course at one of the four institutions, including men and women taking advanced and first-year courses and courses specifically designed for pre-service teachers. The interviews were recorded, transcribed, and coded by tagging each text passage with a code, or label, to capture the main idea. As each distinct idea is tagged with a distinct code, the resulting codebook represents the nature, range and weight of participants' opinions and experiences. We gathered related codes under broader themes and counted the frequency of use of each code or theme as a measure of weight of opinion. As speakers often make several points about a particular issue, the number of observations is generally larger than the number of speakers. The content and frequency of codes was also analyzed by student sub-group, including gender and class type; there were too few minority students in the sample to analyze by ethnicity. The student sample and methods of interview data-gathering and analysis are described in full in Appendix A7.[1]

7.2 Student Learning Gains from Participating in an IBL Mathematics Course

Students were overwhelmingly positive about their experiences in IBL mathematics classes: 89% of all statements about how they did or did not benefit from an IBL course referenced cognitive, affective, and other gains made in an IBL class. "Negative" gains statements (only 8% of all gains-related responses) were not criticisms of the experience, but indications that a particular gain had *not* been made. Another 3% of statements provided qualified, ambivalent, or "mixed" assessments about the *extent* of any gain.

[1] http://www.colorado.edu/eer/research/documents/A7interviewMethods.pdf

The benefits described were of five major types, listed in Table 7.1. Some benefits were particularly significant to men or women (see Table 7.2), or to students in first-year, advanced, or pre-service teaching classes (see Table 7.3). Overall, students report growth in multiple domains important to their personal and professional lives.

Table 7.1: Summary of all Student Gains Reported in Interviews

Gains Categories	Positive obs.	% of all positive obs.	Mixed obs.	Negative obs.
A) Cognitive gains	257	32%	0	22
Mathematics concepts and ideas: Better understanding, knowledge of concepts and ideas; knowing more, better recall. Also transfer of content knowledge to other academic areas, other mathematics courses, life in general				
Thinking skills: Analyzing proofs; finding errors; applying problem-solving skills. Critical thinking; habits of mind; flexibility, creativity, intuition; how to start unfamiliar problems. Also transfer: generally useful thinking or problem-solving skills; use of skills in teaching, other academic or mathematics areas, in life				
B) Changes in understanding the nature of mathematics	54	7%	1	8
Increased understanding of the nature of mathematics: Understanding how mathematical knowledge is built; how mathematicians work, how research is done; seeing the possibility to make a contribution to the field oneself				
Change in conceptualization of mathematics: General changes in ideas about mathematics; mathematics is about proving, not computation; is open-ended, about discovering				
C) Changes in learning	183	23%	3	6
Changes in personal learning: Metacognition, self-reflection; independence; seeing multiple ways to solve problems; deeper learning; persistence; discovering new ways to learn & think				
Learning from others: Seeing how others solve problems; other students as a resource; appreciation of learning from others; group work, collaborating with others as a benefit				
D) Affective gains	179	22%	19	32
Enjoying, liking mathematics more; gains in confidence; increased interest or appreciation; more positive attitude				
E) Communication gains	137	17%	1	7
Writing, presenting proofs; increased ability to explain material to others; ability to organize a presentation; giving or receiving critique. Skills transfer to other areas, to life				
Total all gains	**810**	**100%**	**24**	**75**
Total percentage of all gains-related observations	***89%***	**--**	***3%***	***8%***

Table 7.2: Student Gains by Gender, from Student Interviews (positive reports only)

Gains Categories	Obs. by men (n=29)	*% of all obs. from men*	Obs. by women (n=38)	*% of all obs. from women*	All obs.* (N=68)
A) Cognitive gains	134	*52%*	119	*46%*	**257**
Observations per capita	*4.6*		*3.1*		***3.8***
B) Changes in understanding the nature of mathematics	26	*48%*	25	*46%*	**54**
Observations per capita	*0.9*		*0.7*		***0.8***
C) Changes in learning	88	*48%*	87	*48%*	**183**
Observations per capita	*3.0*		*2.3*		***2.7***
D) Affective gains	84	*47%*	86	*48%*	**179**
Observations per capita	*2.9*		*2.3*		***2.6***
E) Communication gains	77	*56%*	56	*41%*	**137**
Observations per capita	*2.7*		*1.5*		***2.0***
Total all gains	**409**	**50%**	**373**	**46%**	**810**
Observations per capita	*14.1*		*9.8*		***11.9***

* One student did not report his/her gender. A total of 28 observations, distributed across categories, are included in the totals, but not itemized; therefore the row percentages shown do not sum to 100%.

Table 7.3: Student Gains by Class Type, from Student Interviews (positive reports only)

Gains Categories	Obs. by 1st-year students (n=19)	*% of all obs. by 1st-years*	Obs. by advanced students (n=35)	*% of all obs. by advanced students*	Obs. by pre-service students (n=14)	*% of all obs. by pre-service*	All obs. (N=68)
A) Cognitive gains	84	*33%*	125	*49%*	48	*19%*	**257**
Observations per capita	*4.4*		*3.6*		*3.4*		***3.8***
B) Changes in understanding the nature of mathematics	20	*37%*	24	*44%*	10	*19%*	**54**
Observations per capita	*1.1*		*0.7*		*0.7*		***0.8***
C) Changes in learning	50	*27%*	86	*47%*	47	*26%*	**183**
Observations per capita	*2.6*		*2.5*		*3.4*		***2.7***
D) Affective gains	58	*32%*	84	*47%*	37	*21%*	**179**
Observations per capita	*3.1*		*2.4*		*2.6*		***2.6***
E) Communication gains	39	*28%*	77	*56%*	21	*15%*	**137**
Observations per capita	*2.1*		*2.2*		*1.5*		***2.0***
Total all gains	**251**	***31%***	**396**	***49%***	**163**	***20%***	**810**
Observations per capita	***13.2***		***11.3***		***11.6***		***11.9***

In the following sections we describe the types of gains collected in each broad category shown in the tables, and illustrate with quotations from the interviews. Quotations are labeled with coded references to an individual or focus group interview to provide a traceable link to the raw data; but to preserve student anonymity, no other identifiers are given.

7.2.1 Cognitive gains were the largest group of gains reported by IBL students.

Students most commonly described cognitive, or intellectual, gains: growth in their ability to understand or connect ideas across areas of mathematics and in their thinking skills. Such comments comprised nearly one-third of all positive gains observations; on average, students offered almost four comments each about their intellectual development. Men and first-year students offered slightly more comments than others on cognitive gains (Tables 7.2 and 7.3).

Students' observations on cognitive gains divided into two subcategories: *better understanding of mathematical concepts and ideas* and *improved thinking skills*. In both subcategories, students described how such gains also transferred to their academic work or to life, in general.

7.2.1.1 Students gained understanding of mathematical concepts and relationships.

Students clearly detailed how, as a result of taking an IBL mathematics class, they now had a clearer understanding of mathematical concepts and ideas. Students emphasized how working through the material for themselves gave them a stronger and deeper grasp of the material, that they had come to "really understand mathematics." "Fuzzy concepts" became clear because they had unpacked them fully. This meant that they could better recall the material they had learned.

> Once you get it, you really get it. And since you have to discover it for yourself, it makes it so that you actually know it…. With this way [of learning], you can actually go and apply it and really remember it. (G111)

> I feel like I just actually understand where everything comes from, and one thing follows from another better. …I took calculus before, but I definitely didn't—like, I remembered the important parts of it, but I didn't remember all of it quite as well as I would have hoped. And I feel like, after this, I really understand the concepts better, and know how it all fits together and everything, and that's very much deeper learning. (S114)

By working things out for themselves, students made connections among ideas.

> I feel like I am kind of going back to stuff I was learning in high school and actually understanding why I was doing it. I think it really ties everything together—and not just everything in the course, but all of math. (G137)

Though they had covered less material, they knew the material in depth.

> The biggest differences between this and another class is, it seems like we don't cover as much, like we're a section short or something. But what we do cover, we definitely know better than the [the traditionally taught] class. (S126)

And their own investment of time and effort in understanding was central to this deep learning:

> But an IBL class, the student has to really work out proofs, really has to understand the concepts in order to get the math proofs that are demanded of you. And because of the more intensive work that is involved on the student's part, I learned a lot more, and a lot better, in the IBL class. (S115)

7.2.1.2 Students reported improved thinking skills useful in mathematics and beyond.

In describing gains in thinking skills, students reported being able to think more logically, thoroughly and analytically. Their IBL experiences had "built a good foundation" (G124) of problem-solving skills, enhancing their ability to attack a problem and think through it completely and critically.

> It teaches you reasoning skills. We worked a lot on proofs and how to formulate a proof and organize it and stuff, which I think was helpful just in organizing your ideas in general. (S109)

> I'd say it helps my critical thinking skills, just because I need to look at something and figure out how to prove it. (C114)

Students had developed their ability to build rigorous proofs.

> I can get through the proofs. I mean, they're not always perfect. They're not always exactly right. But they're more or less along the lines of correctness, and the trick now is polishing them to make them perfect…. I'm comfortable with just little mistakes, as opposed to before—I'd just be scratching my head because I didn't understand how to write a proof. So it's certainly been huge. (S101)

Some students mentioned developing a sense of mathematical intuition or creativity and gaining new ground in approaching novel and challenging problems. This student also noticed how this had increased his sense of self-efficacy:

> Most things that I see that are math—or even other subjects—are a lot less intimidating if I don't know how to initially do it, because I feel that I can find out by myself without needing to go look up how to do something. (G120)

Some students mentioned how their improved reasoning skills were useful in other courses and in life in general. Recognizing these new skills fostered confidence, too.

> The best thing, probably, for me, is the problem-solving…. Just the ability to look at something that you're not used to and to be, like, not intimidated by it, and to feel like you can solve it, has been helpful in a lot of ways, not only in math, but also in other subjects. ...In physics, when a lecturer shows a derivation, it's easier for me to comprehend and follow. (G120)

> What I love about the Moore Method is that it develops a type of thinking that is applicable to any field…. When you have to prove something, you're given a set of assumptions and using those assumptions, you have to come up with a conclusion. So

> [in] my work, which deals with international relations, it's the same process. I'm given a set of events, and I make a few assumptions, and from that I have to extrapolate extra information. So it's very similar thinking even though, superficially, the fields look very different. (G141)

Pre-service teachers noted ways in which such skills would transfer to their future teaching, helping them "break things down for kids" (G121) and developing a habit of curiosity and skepticism that they saw as beneficial to their future students—sometimes to their surprise.

> A lot of people were telling me, "Yeah, it's about teaching math," so I was like, "Okay, how hard can that be?"…. [This class] definitely twisted my mind about how to teach math…. It's less structured in the sense of, like, processes and procedures. It's definitely thinking outside of the box, working with others, and, like, why things are the way they are rather than, 'They are the way they are. Just accept it.' So this class allows us to question a lot of things, 'Why does this algorithm work?' And we don't know; we were just taught in schools about it. I think, after taking this class, if we were placed in junior high schools and high schools, I would want to teach them this way…. I think students learn way more doing it this way than how we were taught it. (G122)

In sum, students detailed a rich variety of cognitive gains that emphasize their comprehensive understanding of mathematical concepts and learning these at a very deep level. Some students saw how their mathematical knowledge transferred to other coursework. They also credited inquiry-based learning in mathematics with building thinking skills that were highly transferable. Learning how to do proofs through IBL developed their abilities to approach and solve novel and challenging problems, build and refine logical arguments; and fostered more flexible and creative approaches to problem-solving. Such thinking skills transferred to other areas of math, to other academic areas, and to life in general. Pre-service teachers saw how these thinking and problem-solving skills transferred to their future teaching careers.

7.2.2 A smaller number of students reported gains in understanding the nature of mathematics.

Highly related to students' observations on their cognitive gains, were their comments about changes in understanding the nature of mathematics. These observations also divided broadly into two sub-categories: *increased understanding of the nature of mathematics* and *changes in the conceptualization of mathematics*. Though fairly small in number at just 7% of all gains, observations about understanding the nature of mathematics as a discipline mark a higher-order intellectual gain, achieved by some students. Thus, more than gains in knowledge or problem-solving skills, gains in understanding the epistemological nature of mathematics and "what mathematics really is about" mark important shifts in students' development as mathematicians.

7.2.2.1 Some students reported increased understanding of the nature of mathematics.

This group of observations can often be characterized as revelatory, as students gained a sudden insight into the nature of mathematical knowledge. In understanding how mathematical knowledge is built, they saw that they, too, could make a contribution to the field and be a "real mathematician" (G139).

> The best thing, all the theorems up to now, I know that I can prove them, and I have proved them. So I have the confidence that math is something that's not imposed upon me but something that is true, and that I can verify. So I think there is that, I guess, confidence in math. I believe math. (*laughs*) That sounds kind of funny, but I believe math is made true because I can figure it out on my own. (S116)

Again, students often connected their gains in understanding the nature of mathematics to affective gains in interest and confidence.

> It's actually exciting, because you're approaching the material the same way that the mathematicians at the time approached it, where they didn't necessarily have the answers either. We just have the benefit of, we don't chase the false paths and we were led the right direction. But in the same way, we are making these discoveries ourselves. (G139)

7.2.2.2 Some students reported changes in their conceptualization of mathematics.

After taking an IBL mathematics course, some students changed their conceptions about what mathematics is or "how it works." Some realized that mathematics is about "proving," a "completely different beast" (S135) than computation, as one student put it. He elaborated, "Just being able to think in the abstract was something that was new" (S135).

> The one thing I realized after that IBL course was that mathematics is kind of an abstract field. Before, I thought that mathematics was full of formulas. (G139)

A few students came to see that there wasn't just one right answer: there were multiple paths to a valid proof.[2] Often this involved giving up earlier, comfortable notions of mathematics as a clean and simple kind of knowledge.

> I loved math in high school because there was always one answer. And then I got to this class and there's so many ways to do it, and so many ways to interpret it. (S128)

> I probably did come into college as a math major because [I thought] this is straightforward; I don't need to know why. …It's just something that you do these operations and you get the right answer. But I guess throughout the last three and a half years, my views have changed. I think it's really neat how you can come to the same answer two different ways. (G134)

For pre-service teachers, such insights shaped their future approach to teaching mathematics.

> IBL opened my eyes to [the notion] that math isn't necessarily something that's one way only, which is something that I've always been taught. ...It's definitely introduced me to

[2] In the "Changes in Learning" category, we discuss students' appreciation of multiple approaches to problem-solving in terms of personally applying this in their learning. Here it is discussed in terms of a change in their conceptual understanding of mathematics as a discipline.

> the fact that you can approach math from so many ways.... You can show students that [think], 'I'm never going to use this,' that there really are applications for various things, and even things that seem to them they won't use. I think it's important to show them contexts and situations where there really is something that's useful. (S123)

Others realized that mathematics is about "discovery" rather than memorization or rote application of procedures.

> The seminal moment for me was when I figured out on a proof that there was an issue that we had to be careful of. So it was kind of neat to finally see that and say, 'Oh! I can't just blindly do what the problem statement says. I have to consider this little case here.' So there's just this big discovery moment, like, 'Hey! I discovered that myself!' (G139)

Overall, students' comments about changes in understanding the nature of mathematics suggest that they are developing a more realistic view of the discipline as a body of knowledge and ways of knowing that have been built over time. They also gained insight into what it meant to be a mathematician and realized that they, too, could make mathematical discoveries. They moved beyond long-held conceptions to understand that mathematics is not about rote computation, but about a process of proof, and that there are numerous ways to approach and solve a problem. Many of these comments mentioned affective gains, as well—students' enjoyment of this discovery process and of the insights they had gained. Though this category is small, this set of observations reflects students' personal and professional growth as young mathematicians.

7.2.3 Students reported changes in their approach to individual and collaborative learning.

Twenty-three percent of students' positive gains observations described changes in their approach to learning: developing greater insight into their own learning and maturing in their personal behaviors and attitudes toward learning. These insights include both changes in personal approaches to learning, and in learning from others as they became more aware of other students as a learning resource. Overall, students in pre-service teaching courses offered a rather larger number of learning changes observations than did students in math-track IBL classes (see Table 7.3). These students suggested how their learning insights would transfer to teaching and their future students' learning processes.

7.2.3.1 Changes in personal learning included awareness, persistence, and independence.

In this subcategory, students described developing insight into how they learned best and awareness of their personal learning style. Some noticed their increased ability to work independently.

> I definitely feel a little more confident in trying things out on my own first to try to figure something out, rather than just depending on someone completely introducing every idea to me first. I think that was useful. (S133)

> One thing that I've learned is if I wanna understand something, I just need to spend some time alone with it, you know? …I think that once you spend some time alone with it, then talk to other people, that *really* helps solidify it. (G118)

Others described their growing perseverance in solving problems. As one student remarked, "It was frustrating, but I think it was good for me—'cause it teaches you to persevere, and to keep thinking about a problem in a lot of different ways" (S109). Another student described a community-reinforced insistence on getting the homework solved:

> Sometimes we wouldn't get home for dinner until midnight, 'cause we'll be locked in a classroom and we will refuse to do anything else until we solve this one problem. We'll spend like ten hours, everyone cramming, and just sitting there silent sometimes, trying to think how to—like, [could] we visualize this problem in a different way? …You get really persistent. (G104)

Students began to develop "personal ways of getting at problems" (S133) as they saw over and over again how problems could be—even should be—approached from multiple angles. This was particularly powerful for pre-service teachers, who saw it as providing multiple ways to help their future students, and as freeing them from the perception that there was only one right way to do things.

> I did know this, but [I was] continually being reinforced that there are many different ways to come to a problem. And I think probably the biggest gain for me, just because I'm going to be a teacher, is having another way of teaching in my repertoire—so that if I need to use this, I can. If I need to do it the other way, I can do that, too. (G110)

> It just opened my eyes that math isn't necessarily something that's one way, and one way only, which is the way I've always been taught…. It's definitely introduced me to the fact that you can approach math from so many ways, and math doesn't have to be this looming, scary subject for students—us or elementary students. (S123)

Other students felt these new approaches to learning would help them in other classes or non-academic tasks.

7.2.3.2 Students valued the efficiency of collaborative work and insight gained from others.

In the second subcategory, students reflected on how their IBL peers supported their learning. They named clear and practical benefits of working in a group, as each member brought distinct ways of seeing, and thus potentially solving, the problem. As one student succinctly put it, "If you have two people working together, you have at least twice as much ideas as you would have as a single person" (S116).

> The whole class is functioning together. It's kind of like a collective brain working out all the problems. If I don't know something, there's someone in the class that does know and [can] kind of fill in the gaps.... It's really kind of interesting to see how one person's

> viewpoint sometimes makes things a lot more clear than what the professor was originally presenting….. I think that's why it feels easier, is because it's almost like the whole class is collaborating to get through everything, instead of me being accountable for everything on my own. (G106)

As this last quotation highlights, collaborative work was not only more efficient, but more interesting. Students appreciated learning from listening to other students' explanations and seeing how others solved the same problem in a different manner.

> For me, I like when other people present, even if I have the problem completely done. I like to see how they do it. A lot of times [the professor] will ask the people to present…. So he'll call people with kind of interesting ways to look at something,… because it's enriching for other people to see the different approach to the same problem. (G120)

In sum, students described an array of gains related to their learning. Increased awareness of their own learning style could also combine with peers' different approaches to improve problem-solving, individually and collectively. Pre-service students could see how an increased understanding about learning would transfer to their future teaching. IBL experiences in mathematics had also taught them to learn more independently and to persist through a difficult problem. Classmates were a valuable resource that enabled and supported learning.

7.2.4 Affective benefits included enjoyment, confidence, and increased interest.

At 22% of all positive gains observations, the third largest category of gains includes student observations on the affective benefits of an IBL mathematics experience—influences on their confidence, attitudes, emotions or beliefs. On average, students offered 2.6 positive comments each on this type of benefits (Table 7.2).

Most commonly, students noted enjoying their math class. "If there's one class out of the week I really enjoy going to, it's this one!" enthused one student (S101). Describing class work as "hands-on" and "cool," another student clearly linked this to her motivation to attend class and work hard: "Since I enjoy it, I feel more motivated to go" (S107).

> The *best* thing is just that it's really *fun*! It highlights how enjoyable math is…. My long-term plans don't really have much to do with math; I'm just sort of doing it now 'cause I really like it. And so, IBL has made me continue to like it and…I guess (*pauses*) I'm gonna look for classes like this, in math, until I graduate. …I don't really wanna go back to regular style, because I know how good this one can be. (G118)

Students also described important gains in their confidence to do mathematics and to solve mathematical problems. "Even though I feel like I wasn't really as smart as a lot of people in my class, I realized that if I worked really hard, I could do just as well as them," said one student (S109). They felt more confident that they could "figure it out themselves" and gained a sense of achievement from having learned how to construct proofs.

> I learned a lot more than I do in a lecture class. I learned why things are—and you remember things a lot more because you had to struggle through it and you feel really good at the end of it. Like you accomplished a lot. When [the course] was over, I felt awesome. It was a pain getting through it, but then you feel really accomplished, like you actually did a lot of stuff on your own. (G138)

They took pride in the "explicit and strong" quality of their proofs, the "flow" of the argument, and the quality of the work evidenced in journals that documented the work they had done.

> You sort of develop a pride when you're done with your journal. It's, you know, like twenty pages typed up, and it *really* feels like you've accomplished something. (G118; emphasis in original)

They felt more at ease in presenting proofs in front of a class, comparing themselves to people who take "normal classes" and do not develop comfort in presenting:

> This [IBL class] does make it easier to get up in front of someone, show them what you did, and feel more comfortable in defending what you did. (S126)

Pre-service teaching students noted greater confidence in their ability as future teachers:

> I think it's made me a little more excited actually to teach math. Before I was like, 'Oh, this is going to be kind of a big headache, because a lot of kids don't like math.' But when you take it with this approach, it actually gets them involved, especially if you pick subjects that are very tangible to the kids…. Then they'll get excited about it and be willing to share how they solved it, and they'll be more proud of how they figured it out rather than just saying, 'Yes, I understand the algorithm, let's move on.' (G121)

Several students found IBL classes more interesting. For some, this was simply a better way to get through class: "It's a lot more interesting than sitting and listening to a lecture" (G120). For others, this was sufficient to extend or renew interest in mathematics as a field of study:

> It actually got me interested in mathematics. Previously, math courses were all kind of like an exercise in jumping through hoops, or so it *felt*. And when I took number theory, it felt like I was actually *learning* something that was actually *adding* to my understanding of math and logic. (S132; emphasis in original)

> It certainly renewed my enjoyment for math…. Since I was really young, I loved mathematics and played with calculators and stuff, but I have to say that, after taking three and half years of math and… just hoping to get through the last few semesters, I'd have to say this has renewed my excitement. (S101)

Among both mixed and negative gains statements, the largest group had to do with affective gains. A majority of these were statements about not enjoying the IBL approach. In these comments, students often acknowledged benefits to their learning, but did not like particular aspects such as group work, the workload, or feeling frustrated when they could not solve a problem. Since negative and mixed comments account for 22% of all affective gains statements, this provides a rough estimate of the relatively small fraction of students who, at the end of an IBL experience, still do not prefer it whole-heartedly.

In sum, students noted a number of affective gains from participating in an IBL mathematics class. They enjoyed mathematics more and found it more interesting as a subject. Perhaps more importantly, they developed confidence in their ability to do mathematics. Indeed, many expressed pride and accomplishment at having struggled through a proof. A few students noted confidence in presenting their mathematical thinking. Pre-service teachers gained confidence that they could teach mathematics well. Finally, a few students noted a more positive attitude towards, mathematics as a result of taking an IBL mathematics class.

7.2.5 Students gained skills in communicating mathematics.

The final benefits category gathers students' observations on gains in communication skills; these were 17% of all positive gains comments (Table 7.1). Men offered nearly twice as many observations on communication skills than women (Table 7.2).

Students detailed how taking an IBL mathematics class improved their ability to write and present proofs—a gain that they expected would help them in future math classes.

> When I'm writing [a proof], I started out having a tendency to leave out some smaller details. And occasionally that would lead me to actually make a small assumption that was wrong, and then that would turn into like a big deal. It's definitely taught me to be more careful…. (G140)

> I definitely have learned to write proofs better than I would if I was just in a normal math class. And I know that in higher math classes, it's a lot of proof writing. So this is just a good background from that point… just all the practice of writing proofs and everything. (S114)

Communication gains were seen generally beneficial in improving public speaking skills, particularly for students who planned to go on to teach one day. Students also recognized their increased ability to explain the material to others. Those who intended to teach especially appreciated the importance of communicating their thought processes effectively:

> Having someone talk me through it also helps me understand what they're doing and how their thought process is working. And I think it's an important skill, to be able to talk through what you know, how you thought about a problem, and how you got to where you did. (G141)

> I'm definitely going to have my students do some type of presentation. …I think it's important, because I found that your ability to communicate and explain your logical steps and your thinking through a problem really helps you as a learner. (G138)

A few students described gains in being able to "stand up to critique," "defend yourself" and respond to comments on their feet. Again, students saw ways in which these gains would transfer to other areas and to life in general:

> The fact that you have to be able to present your work and defend it and have other people critique it, that's valuable. And that's not just in mathematics; that is just a life skill, generally. (G139)
>
> Every time you solve a problem, you have to write in a fashion where somebody who has no knowledge of the class can understand it. And I think that's a really good advantage, and not even just for the class, but for anything else that you may need to do. (G104)

To summarize, students described a variety of gains in communication skills. Central to their mathematical work was learning to write and present proofs. Having to carefully explain their logic increased their ability to communicate ideas effectively to others while simultaneously solidifying their own understanding. Some saw learning to give and receive critique and to defend their answers on the spot as benefits. Pre-service teachers recognized ways in which communication skills would transfer to their teaching or to life, more generally.

7.3 Learning Processes Reported by IBL Students

While students' reports of their learning gains were significant within the interview data, they also described in some detail the processes that they saw as leading to their learning. With nearly 1200 distinct observations comprising some 40% of the total student codebook, these data offer great insight into how students learn in an IBL classroom. Here we highlight key elements of the processes that students described as important to their learning of mathematics.

7.3.1 Deep engagement with mathematics and peer collaboration were the twin pillars of learning in inquiry-based mathematics classrooms.

Two broad elements emerge from student interviews as central to the process of learning in an IBL mathematics classroom: deep mental engagement with mathematics, and peer-to-peer collaboration. These two ideas were so common, and so intertwined, that we refer to them as "twin pillars" of IBL instruction: neither alone is sufficient to explain the learning outcomes, but together they account for much of what students said was beneficial about their IBL experience. Students saw both engagement and collaboration as integral to the IBL approach. While they also discussed technical aspects of the course—such as course materials, assignments and assessments, and access to instructors' help—and characteristics such as the course pace and classroom atmosphere—these appeared to function as auxiliary to the twin pillars, by aiding, abetting, or hindering students' engagement with the course material and the effectiveness of their collaborative work. In this way, instructors' decisions about these course features can be critical to the success of an IBL course, as we discuss in Chapter 8.

7.3.1.1 Deep engagement fostered deep understanding and relied on students' motivation and effort.

Engaging with the mathematics was built into the IBL approach, as one student put it: "Students aren't just copying from the board and learning just algorithms, [but] learning how things work" (S108). Another explained how this emphasis on "how things work" led to deeper learning:

> I think I understood it a lot more, because I think the lecture-style classes don't really explain why you are doing this or how this came about. But the IBL classes are the ones that you know why, because you figured out why, and you know how it came about because you made it come about. So did the whole class. But *you* figured out why the negative numbers are negative numbers. (G137, emphasis in original)

Such deep learning was by no means automatic within the IBL approach, but depended substantially on individual effort. Students made 96 comments about the effort they expended in an IBL course; over two-thirds of these observations addressed the higher effort and greater time input they saw as required, including time working alone, working with others informally, and attending office hours. While students' learning may thus be due in part to greater time on task, they were motivated to expend this effort because the tasks they were assigned seemed meaningful to them—not just "busy work," as this student explained.

> Student: I like it a lot more. I feel like I just actually understand where everything comes from, and one thing follows from another better. It's like no memorization or anything like that, and of course there is much less computations, so it seems like no busy work or anything like that.
>
> Interviewer: So you don't feel that you have had to work more, work harder?
>
> Student: Well, I do have to work harder, but I feel like when I've finished a proof or something I've actually accomplished something. Unlike just solving—multiplying two things or something like that, or integrating a function—I've actually like proved something, which I like. (S114)

Even as they described hard work and high effort, many students articulated this tradeoff as worthwhile: "I know that I will get a lot out of what I put into this class" (S116). To succeed, however, students saw it as necessary to be "motivated," to "take initiative," to be "plugged in" instead of "out of it." Indeed, instructors noted that students lacking motivation were the most likely to slip through the cracks. However, while motivation and effort were to some extent pre-conditions for success in an IBL course, they also appeared as outcomes of the method, as students were motivated to spend effort on challenging problems that interested them.

> I spend more time on homework for IBL classes. But it's not like, 'Oh my, I have to do that homework tonight.' It's like when I get home, that's usually the first thing I work on—so I feel like the homework isn't as much of a chore as other homework. Some other classes' homework, where you are just looking at your notes and then just doing the exact same thing, feels like busy work. Whereas with the IBL proofs, ...I'm actually going to learn at the end and have something to take away from this homework, instead of just getting a grade and copying the answers from the back of the book. (G137)

As another sign of engagement, students described experiences of "getting in the zone" while working on proofs: losing track of the clock, having sudden revelations in the shower or the gym, spending many hours on a problem for pleasure or out of determination to solve it. After a late and frustrating night of failing to crack a problem, one student described a morning aha:

> I would—seriously—be in bed and my alarm would go off and I would just be laying there in bed and I would be like grabbing my papers to write down my proof, because it just came into my head. (G138)

In sum, deep engagement with the mathematics is one important source of the positive learning outcomes students reported. As these quotations illustrate, students are motivated by solving interesting, non-repetitive problems. As they see the payoff of their effort in the form of significant learning, they are further motivated to work to understand the material.

7.3.1.2 Peer collaboration made IBL classes enjoyable, fostered confidence, and required communication that developed skills and deepened understanding.

Even more important in students' reports of how they learned was the role of collaboration. Students described three types of collaborative work: structured small group work in class, informal small group work when studying out of class, and whole class discussion. From observations (Ch. 2), we know that structured small group work was extensively used in some IBL courses, such as what this student describes:

> We break out into groups before we present the problems, and it helps because if you're struggling with it other people can help you. And then the teacher is also available to come by and aid our group, if we're really having difficulty with something. (S101)

Whole-class interaction was common across all IBL courses. It most often took the form of class discussion and critique of a solution presented on the board:

> Like, the professors would go, 'Why don't you go put it up on the board?' And you write out your explanation and then you say it to other people and then they tell you, 'Well, I think you could've done this differently in your proof.' ...And, 'Here's where you're lacking.' So, it would be really class-based. (G111)

Finally, students reported out-of-class informal group work to be common and even essential to succeeding in an IBL course.

> Everyone would see a problem differently, and so when you have so many different ideas coming together, it's a lot easier to think of a solution. So I think we sort of rely on each other, because I really… I could not have done any of those homeworks by myself… but I contributed my own portion too. (S109)

Although students described different forms of group work, they saw them as highly related and generally did not distinguish the benefits gained as related to one form of group work or another. Rather, they saw group work as intrinsic to the IBL approach.

> You end up working with people *in* class and *out* of class, and it's the same sort of environment. Whereas in a regular class you would work on your problem sets [together]

> and that would be fun, but then you'd go into class and you wouldn't be collaborative anymore. So there's kind of a nice continuity to how you do all your work in IBL. (G118, emphasis in original)

Students cited many ways in which working with peers benefited their learning. Many referenced the collective power of multiple brains applied to a problem: "If I don't know how to do a proof, probably somebody else does—in which case we can all learn how to do it" (S114). As the same student noted, this also led him to appreciate that there could be multiple ways to reach a solution:

> It's good to see how other people do [the proofs] too, just to look at the different ways of thinking of things, and realizing that there are different strategies you could take, and that occur to some people and not others. (S114)

They cited the efficiency of group work in generating ideas and assessing the likely fruitfulness of particular problem-solving strategies:

> If you have two people working together, you have at least twice as much ideas as you would have as a single person. And also another benefit is that—there are proofs there are so many ways to go about, so if you choose one way and if that way was wrong, you waste so much time. But if there is another person who can give you feedback—if I would say, 'Let's do a proof maybe by contradiction,' and if I were on my own, I would just do it. But because he is there, he would tell me whether that's a good idea or not. And by doing that we could, I think, save a lot of time. (S116)

Because group work required them to explain their own ideas and understand others', students found that their mathematical thinking became more clear in the process. Students also linked group work to their gains in communication skills, which they saw as highly applicable to other classes and to their future professions.

> We all have our different approaches, and then we have to explain it to the entire class how we got to that answer. I think it's helpful for me, just in terms of explaining things to others, but also in terms of being a teacher. You'll explain the way you want something to be done, or if a student's doing something a certain way, explain how they're doing it. So I think it's really helpful in terms of that in preparing me for how to explain my thought process in classroom. (S123)

Finally, students found group work enjoyable. Class time was interesting and even fun; the high level of student interaction established a positive atmosphere where it was safe to share ideas.

> I like the group work. And it's not intimidating when you're able to work in groups and you're able to come to these conclusions together, and then present them to the [rest of the] class…. It's cool, and it's definitely more fun than just having someone throw answers at you or problems at you. (S108)

Interestingly, students' emphasis on collaboration contrasts with the views of instructors, who emphasized engagement more strongly. This difference in perspective makes sense: students pay more attention to their peers than instructors do. Compared to their other classes, the collaborative atmosphere is distinctive and thus striking to students. And students report outcomes such as fun and friendship, which matter to them, beyond the intellectual outcomes in which instructors take more interest. Students also collaborate outside of class, which is less observed by instructors but is important in students' accounts.

7.3.1.3 Collaboration and engagement interacted in a mutually supportive manner.

Peer collaboration and deep engagement are "twin pillars" of IBL because both are important processes for student learning, but they are also "twined pillars" that interact in multiple, mutually supportive ways. When students engaged deeply with the mathematics on their own, they came to class well-equipped to contribute meaningfully to group work. In communicating their ideas and listening to others, they deepened and crystallized their own knowledge.

> When you go away, and you think about it with your peers and stuff, then you're able to like pull it apart and really see what the teacher means. And that's where you claim it for yourself. (S119)
>
> One thing that I've learned if I wanna understand something, I just need to spend some time alone with it, you know. So I think that that's something that I've just realized through IBL. And I think that once you spend some time alone with it, then talking to other people *really* helps solidify it. (G118, emphasis in original)

Another student connected her understanding of course material with the necessity to explain to others: "I think the group work is really a tool for us to talk through our ideas and have someone else there to analyze them while you're doing it, or help you." The process of presentation and critique allowed students to dig deeper into the material and refined their critical skills, as two students described in this dialogue.

> *Student 1:* It's like students [are] teaching the class—well, not teaching the class, but like, we present our ideas, as opposed to someone standing at the front and saying, 'So this is right.' Not just the facts presented—it's more like, students will present what we find and then we all have a chance to question what was presented and debate whether it's right or not.
>
> *Student 2*: It feels like the students as a whole, we critique the proofs that are presented. Whereas in a lecture-based class, you just sit there and write down everything the professor says and get up the nerve to ask a question if you are confused about something. With this, we go very deep into the subtle parts of each proof and what needs to be in a proof and what is maybe unnecessary. (G134)

Moreover, class discussion gave students opportunities to get feedback on their ideas, thereby developing confidence that they understood the material and were ready to move on.

> I think the most beneficial thing we do in that class, after we discuss the problems in groups and try and figure them out, then she asks us to volunteer to go up to the board and explain it…. And by doing that, it allows us to maybe crack any misconceptions, it allows her the chance to correct misconceptions, so we're actually learning the material. And although we took some time to figure it out on our own first, we're not coming away with these incorrect ideas. (S108)

Finally, collaborative work motivated students to prepare for class. Depending on others to get through the material made them feel accountable to their classmates and fostered a desire to contribute meaningfully themselves. As one student said, their work was "independent from the professor, but it wasn't independent from each other." (G103)

> I've noticed with myself, now that I'm in this group setting, it's like I want to figure it out, I want to be able to do it on my own. And then I kind of reach out to the group if I don't get it, or if I do have an idea. And we're able to support each other, rather than, like, just being alone. (*laughs*) (S125)

In these ways, individual engagement and peer collaboration work together to support students' learning of mathematical ideas, development of problem-solving skills, and the affective outcomes that are so important in motivating them to invest time and effort in the course.

7.3.2 Further analyses will analyze specific course components and student difficulties.

The student interview data on learning processes is very rich, meriting more analysis than can be incorporated into this initial report. Future analyses will address students' observations about specific aspects of their courses, including:

- General course features, such as the overall course pace, workload, and learning goals
- Specific structural elements, such as materials, assignments, assessments and grading
- Day-to-day interactions, such as peer dynamics, the classroom atmosphere, and the availability and nature of help from faculty and TAs
- Individual student learning processes, such as problem-solving, effort, and metacognition, or thinking about their own thinking
- The nature and sources of student resistance to inquiry approaches.

Many students offered thoughtful critiques of the IBL method, including ideas about what was "best" and "worst" about their course. Students' reports of difficulties are especially helpful in providing concrete advice to instructors about how to refine their teaching methods. Among student-reported difficulties, the largest group of observations identifies issues of implementation —not problems fundamental to the IBL approach, but aspects of specific courses that could be revised and improved. A second cluster of observations addresses the support that students need to succeed in an IBL course, especially their needs for signposting of where the course is headed and why, and for scaffolding to gradually develop the required skills rather than being "thrown into the deep end." Only a small fraction of student comments suggest genuine resistance to the approach, suggesting that—when it is well-executed—most students are amenable to an IBL experience.

7.4 Summary, Strengths and Limitations of the Interview Study

In interviews, students detailed a rich and varied range of benefits from IBL mathematics courses. These experiences contributed to several areas of their personal and professional growth: gains in their ability to understand and connect mathematical ideas and growth in their thinking skills; increased understanding of mathematics as a domain of knowledge and of how they learn; enjoyment and confidence in their ability to do mathematics; and gains in communication skills. While students commented most on their intellectual gains, the other learning gains categories are also substantial. The interrelation of benefits is particularly evident in the ways in which students connected new understandings of mathematics concepts, the nature of mathematics, and their own learning to their enjoyment, confidence and interest.

In analyzing gains by student sub-group, we did not find many differences. Men offered more observations, in general, but there are no differences in the types of gains emphasized by men or women. This may seem somewhat puzzling given the strong gender differences seen in the survey data (Ch. 3). But the survey data suggest that IBL courses level the playing field for women: women and men in IBL courses report comparable learning gains, while women in traditional courses report lower gains than their peers. Therefore, the gender difference in experience does not lie in the IBL course, but in the non-IBL course, where women's experiences are more often discouraging. Thus the interview data corroborate the survey findings in suggesting that the IBL classroom experience is equitable; both men and women have positive experiences that support their growth in all reported areas.

The variations by class type are more pronounced. On a per-capita basis, first-year students reported notably more gains in understanding the nature of mathematics and slightly more cognitive gains than did students in advanced math-track courses or those designed for pre-service teachers. This may reflect the novelty of engaging in an IBL mathematics course for the first time and the contrast with their experiences in high school. Pre-service teachers found IBL particularly beneficial in fostering changes in learning, compared with students taking math-track courses. This finding makes sense given that students in pre-service IBL mathematics courses are interested in learning, generally, as future teachers. Among pre-service elementary teachers in particular, prior mathematics learning experiences may have been negative; therefore the IBL experience may be unexpectedly pleasant and productive, giving them more reason to reconsider their approach to learning mathematics. Pre-service teachers' lower emphasis on communication gains may reflect the lower emphasis on formal proof writing in these courses, relative to the math-track courses. These were the only apparent variations by sub-group.

Student accounts of their learning processes emphasize the dual importance of deep individual engagement with mathematics and collaborative work with peers. These two elements mutually reinforce each other in several ways. For those familiar with Moore's original method, the importance of deep engagement is not surprising—but the crucial role of collaboration may be surprising. The significance of collaborative work reflects both deliberate changes by instructors to explicitly incorporate collaboration, and students' own pursuit of opportunities to work with and learn from one another. Our finding here that collaboration contributes importantly to student learning is consistent with modern educational theory and empirical work that emphasizes socio-cultural aspects of learning.

In considering the strengths and limitations of these findings, we note that the student interview sample is large for a qualitative study, but may nonetheless fail to represent some student viewpoints. The interview sample is adequate for three of the campuses, but the IBL methods characteristic of the fourth campus may be underrepresented in the data because we could not directly recruit student interviewees (see Appendix A7 for details). More generally, since students' participation in an interview was voluntary, students who had a poor or unexceptional IBL experience may be underrepresented, and students who wanted to share their positive feelings may have been more likely to volunteer.[3] Due to the low participation of students of color in the courses available, the sub-sample of students of color is too small for us to separately analyze their gains or experiences.

Sampling limitations like these are most likely to affect what may be missing from any particular data set; they do not affect the validity of what is present. The patterns in this data set are strong. Overall we conclude that the IBL courses that these students experienced had a very clear and positive influence on their learning of mathematics, their skills in thinking and communicating, and their maturation into adult learners.

[3] We have observed that people with highly negative experiences will volunteer for interviews if they feel that this will help fix a problem. People with mediocre or mixed experiences appear to be least likely to volunteer.

Chapter 8: Findings from Instructor Interviews

8.1 Introduction

Interviews with instructors complement classroom observation and student interviews by providing instructors' perspectives on student outcomes, classroom learning and teaching processes, and their own experiences. We sought to address the following research questions:

- What student gains (or failures to make gains) do instructors observe as a result of IBL instruction?
- How do instructors describe their classroom practices, with what rationale(s) behind them? What teaching issues do IBL instructors face, especially issues that differ in nature or scale from those in non-IBL classrooms, and how do they resolve them? What advice do they offer to other IBL instructors?
- What are the costs and benefits to instructors of teaching with IBL methods? What professional resources are needed to begin using IBL methods, and how (if at all) do instructors obtain them?
- What are the personal, professional, departmental and institutional conditions under which IBL courses are effective and sustainable, or not?

The study methods are detailed in Appendix A7.[1] Briefly, in-person or telephone interviews were recorded, transcribed and coded to keep track of what people said and how frequently they said it. This collective report of experiences, perspectives, and beliefs was analyzed for patterns and themes. Frequency counts indicate the relative weight of opinion about particular issues. Because respondents contribute ideas of their own, open-ended interviews are exploratory and can uncover items that the researcher might not think to put on a survey. We interviewed 43 instructors, 23 with faculty appointments (including postdocs) and 20 whose IBL work occurred while they were graduate students.

By identifying the processes by which certain outcomes are achieved, interviews can help to explain the results of other studies. In particular, the instructor interviews characterize the teaching and learning practices in use and thus help us to understand students' classroom experience. They also document the benefits and challenges of teaching this way, the support that instructors need, and the role of IBL teaching in the professional growth of instructors as teachers. In the discussion below, we consider several of these broad themes and identify the patterns and variations in instructors' perspectives. Quotations are labeled with a traceable link to the raw interview data; but to preserve instructor anonymity, no other identifiers are given. Pseudonyms are used when any name is mentioned. We begin with instructors' observations of student learning, then their descriptions of teaching. We conclude with some observations on the costs and benefits to instructors themselves, and on contextual influences on their teaching work.

---

[1] http://www.colorado.edu/eer/research/documents/A7interviewMethods.pdf

8.2 Student Learning Outcomes and Learning Processes: Instructor Perspectives

Instructors' observations importantly corroborate many of the student gains noted in Chapters 3 and 7. Instructors also confirm the important learning processes that students highlighted.

8.2.1 Instructors observed many of the same gains that students reported.

Instructors made a total of 401 observations about student learning outcomes from IBL mathematics courses, of which 336 (84%) were "positive" statements about gains that they directly observed or believed students to make in their courses. In 65 "negative" observations, instructors commented on gains that they did *not* see students making, especially in comparison with traditionally taught courses. Table 8.1 compares instructors' observations on student learning gains with students' own observations, following the same broad coding categories as used for the student interview data (Chapter 7) and survey write-in comments (Chapter 3). The types of gains that instructors observed match well with those reported by students, but the distribution of gains differs in ways that make sense based on differences in students' and instructors' points of view.

Table 8.1: Student Gains from IBL Mathematics Courses as Reported by Instructors

Gains	Positive gains: number of instr. obs.	Negative gains: number of instr. obs.	% of *positive* instructor obs.	% of *positive* student obs. (from Ch. 7)
A) Cognitive gains	98	44	30%	32%
B) Changes in understanding the nature of mathematics	47	0	14%	7%
C) Changes in learning	27	5	8%	23%
D) Affective gains	67	3	20%	22%
E) Communication gains	79	10	24%	17%
F) General & miscellaneous gains; influences on career or education plans	18	3	2%	—
Total all gains	336	65	100%	100%

8.2.2 Instructors highlighted student gains in thinking skills and deep learning of mathematical ideas.

Instructors had the most to say—both positive and negative—about students' cognitive gains (category A in Table 8.1). It is perhaps reassuring that these gains surfaced frequently in both student and instructor reports, as growth in mathematical thinking and understanding is a commonly shared goal of college mathematics courses. Gains in thinking skills were especially important to instructors, who saw students becoming "good math thinkers" (A504) who could approach a problem "like a mathematician would" (A507). Instructors named many of the same problem-solving skills as did students: learning how to write proofs, how to dive into an unfamiliar problem, how to structure an argument, how to spot for themselves when an argument was complete or valid, how to "piece together" information to solve a problem (A501). Others saw growth in mathematical habits of mind: intuition for how to tackle a problem; creativity in finding multiple or novel approaches; a more critical attitude in "asking each other why" (F701).

> When you give students problems just freely to work on, they come up with their own ways of doing things—and a lot of the time, they're unexpected. The techniques that students will master will not always be the ones that you expect them to master, and they will frequently surprise you by coming up with techniques that you hadn't even thought of. That's one of my favorite things about it, is it invites the students to be inventive.... They really do come up with things. (A513)

Instructors also emphasized that IBL led to deeper learning. Having to "digest ideas completely" (F711), they thought, meant that students understood concepts thoroughly and would remember them. As one TA put it, "Because we need these tools, they develop them on their own and then they understand them more" (A515).

However, nearly as often as they noted deep learning, instructors said that IBL students covered less ground. Indeed, the largest group of negative observations on cognitive gains expressed that students learned less material because of the slower pace of IBL courses; nearly all described this as a tradeoff of learning in depth against breadth of coverage. Half of the coverage comments came from TAs at one campus, suggesting that they had discussed this as a group.

Smaller numbers of negative observations noted lower gains in problem-solving when students (especially pre-service teachers) could not give up old algorithmic approaches; lack of exposure to disciplinary standards of proofs (because students were inventing their own ways); and limited depth of insight because the professor could not share his own expertise on the material. A few observed that IBL calculus students in particular lacked exposure to computational methods and examples that might be important to them later.

8.2.3 Instructors noticed students' growth in confidence and enjoyment of IBL classes.

Prominent among the affective gains (category D) that instructors observed was students' growth in confidence: in their ability to understand difficult material and solve challenging problems, to articulate ideas with "the courage of their convictions" (A511), to "do more than they think they are capable of" (A501). Among pre-service teachers, this translated as lessened "fear of math" (A611). As students became more confident in presenting work at the board, this bled over into other behaviors:

> ...I noticed that they are much more likely now to talk in an ordinary class and ask questions and discuss it. Then when they're totally confused, an ordinary classmate, one of them, will make a good suggestion which will resolve the problem at hand. (F706)

Instructors also observed that students enjoyed class, particularly group work, and the feeling of mastery gained from solving a hard problem. Some suggested that students found it exciting to create the mathematical ideas themselves, instead of being given the ideas in a textbook—to discover "the joy of thinking" (F718). In general, instructors noticed that students found the IBL approach interesting and motivating, although a few worried that the everyday routine of presenting and critiquing became boring after some weeks. While faculty are not always aware of students' affective gains (Laursen et al., 2010), which are personal and internal, IBL faculty and TAS alike seemed to recognize their students' affective outcomes from the course, reporting these gains in sizable proportions comparable to what students themselves reported.

8.2.4 Instructors were alert to student gains in communication skills and in understanding the nature of mathematics, but less aware of changes in their approach to learning.

Cognitive and affective gains account for over half of all gains reported by both students and instructors, and were described similarly by the two groups. In the other three gains areas, instructors' and students' perceptions of student gains from IBL courses differed noticeably. Instructors more often than students reported gains in communication skills (E in Table 8.1), and in understanding the nature of mathematics (B), while they less often reported changes in students' approach to learning (C). All three gains areas are especially interesting because they identify learning outcomes that are important to IBL instructors but not universally addressed in college mathematics courses—"stuff that just doesn't happen in a regular curriculum," as one noted (A505). Differences in student and instructor perspectives are also interesting to consider.

As professional mathematicians, instructors had a good sense of what is required to communicate mathematics effectively: using precise language, attending to details, ensuring that all the logical steps are present, writing complete sentences or statements, seeking elegance and efficiency. They observed students improving in their ability to do these things in written work and at the board, as well as coming to understand that these qualities would be expected of their proofs. As one instructor put it, "They are more clear about what they say. And also when they say something, they know that they are expected to explain their reasoning. (*laughs*) So they are very careful about what they say!" (A510) Another instructor connected improved written work to oral presentation: "I think it's easier for them to see if it makes sense or is convincing when they actually try to convince a class of students, rather than just writing it down on paper and turning it in" (F704). While instructors set high standards for mathematical communication, they could clearly detect progress toward meeting those standards.

> Even the best don't have quite the big picture, in the way that I can sketch a proof in three sentences and then tell you the detail. But I think just saying, 'Okay, I have an argument in mind. How can I organize it in order clearly and logically and then present that?'—certainly [that] is something that they have, and something that you could call communication skills. (A507)

Instructors also saw students learning to give and receive critique: to "take a chance and expose themselves to having an argument with somebody" (F711), to "feel a little bit more free to express their ideas, and realizing that saying the wrong thing occasionally is the way to learn things" (A511). When instructors reported gaps in communication gains, these most often had to do with students' skills in reading mathematics and in learning mathematical notation and proof styles that would be better modeled by a textbook than by inexpert peers.

Likewise, instructors were better able to see changes in students' understanding of the nature of mathematics (category B) than were students—again because, as professionals, they had deeper knowledge of just what this meant. Especially in transition to proof courses, they noticed students coming to understand mathematics as abstract rather than computational; as based on proving rather than procedures; as allowing and valuing multiple "right ways" to a solution. These observations are similar to what students themselves reported. However, instructors also took note when students arrived at deeper insights, such as understanding that mathematics is a creative human endeavor involving investigation and experimentation. Using the power of the human mind, therefore, mathematical ideas were accessible to all: "There's a real right answer out there and you can be sure that your answer is right even if someone with authority hasn't told you," said one (A514).

> I think they're learning a lot more about how it works for mathematicians than if it were a lecture-based class where everything's presented as, 'This is the theorem and this is how you prove it,' and it's like the proof is handed down from on high, you know. And math works sort of the opposite way—you have to figure out why things are true. (A509)

While they saw students' understanding of the nature of mathematical knowledge deepening, instructors were quite clear to distinguish what went on in an IBL class from real research, describing IBL as a "stepping stone," a "simulation" or a "ramp" to real research. As one TA elaborated, "They're not discovering new math, right? They're not solving questions that no one else has solved before—but they're solving questions that no one in the class has solved before" (A512). Nonetheless, some felt, IBL courses gave students a glimpse into "math as practiced by real mathematicians, which is itself a fun, enjoyable, creative, artistic act" (A507).

Finally, instructors reported some changes in students' approach to learning (C in Table 8.1), which match students' reports of greater independence, persistence, and ability to work with others. "They now feel empowered to learn things, not just quote them," said one TA (A517). "They really want to think about these things for themselves, and they really spend time thinking about things," said another (A509). Overall, however, learning changes were rather less prominent than in students' reports—understandably so, since these are quite personal changes within students' own minds, observable by instructors only as secondary effects or if discussed with students. We observed that instructors very seldom asked students to reflect on their learning in ways that might help them recognize and make good use of these changes. Students came to these realizations on their own, if they did at all. Overall, however, instructors' observations strongly corroborate the content of student-reported gains, while their differing "yardsticks" for assessing gains account for differences in each group's emphasis.

8.2.5 Instructors perceived that IBL benefited certain groups of students in particular.

Instructors were asked whom they saw benefiting most from IBL courses. A total of 27 instructors offered 74 different theories about differential benefits to certain student groups. The most common type of theory, discussed in 34 observations by 14 instructors, related to students' prior achievement. The level of consensus among these observations was surprisingly high: "weak" students do not benefit or do not succeed (16 observations); "good" or medium students benefit in particular (10 observations); and the "best" students benefit or at least are not harmed (8 observations). One TA laid out a rationale that was representative of others' reasoning as well:

> For the best students it probably didn't really matter that much, in a certain sense. They were very strong and they understood the material. And maybe they approached it in a different way, you know—they were doing a lot of work in the discussion—but they probably would be similarly involved for a non-IBL course. Just for the stronger students.
>
> (*continuing*) I felt it made the biggest difference for the middle section of the bell curve or something, a lot of these students who would not have had to have been involved. They could have just listened to the professor and then gone and done the homework, and then maybe gone to some study group or some tutorial session, and worked through their homework, but kind of being led by the stronger students. The IBL method kind of forced them to take control a little bit more, in terms of putting things in their own words. Putting things not just in their own words, but in their own ideas. So I felt like it boosted the middle range of students.
>
> (*continuing)* At the same time, I thought it was similarly kind of worse for the very lowest students.... I felt like some of them were just kind of lost all year; they were never really quite able to get into the mode so this would work. So every time that it was their turn to present something in class, they would kind of fumble, and would have to say, you know, "Can anyone else do it?" And they weren't ever really able to take part in discussion that way. (G506)

Another common type of theory, offered by ten instructors, related to students' personalities: students who were outgoing or verbal benefited more than students who were shy or quiet, because they were comfortable presenting solutions and participating in discussion—activities seen as integral to achieving IBL's good outcomes. Seven theorized about students' motivation and interests, suggesting that IBL worked best when students were committed to studying mathematics (e.g. math majors); or that IBL was good for students preparing to be teachers.

Five instructors offered a theory about differential benefits to students who are often under-served in college mathematics: women, students from minority groups underrepresented in mathematics, and students from working-class backgrounds. All of these proposed that any particular benefits of IBL to these groups were due to the confidence gains that IBL approaches engendered. Five instructors observed no patterns of difference in benefit to students. Overall, some of instructors' theories about who benefits most were corroborated by survey (Ch. 3), test (Ch. 5), and academic records data (Ch. 6) and thus may help to explain these findings.

8.2.6 Instructors identified deep engagement with the mathematics and collaboration as central processes in student learning.

In describing the learning processes important in IBL classrooms, instructors identified similar elements to those named by students. "I think that the first point of the IBL is to get people to figure out things for themselves and therefore really figure them out at all, really understand them in the first place," said one experienced TA (A514). From the perspectives of instructors, both deep engagement with the mathematics and collaborative work with others facilitated this "figuring out" process.

Instructors emphasized the importance of students engaging directly with the material: preparing for class, coming to class, participating in class, "interacting with the math" (A515). As the largest category of instructor commentary on learning processes, deep engagement was seen as fundamental to student learning of any material: "Anything that you really engage with, you get a lot out of. And if you're just going through the motions, trying to get it done, you're not going to get as much out of it" (A611). In contrast, noted another TA wryly, "The students whose plan it is to only be awake when it's dark out—this is really going to fail for [them]" (A505).

Students' effort was seen as important—necessary to learn the mathematics, but also a life lesson that real learning requires hard work. Like the professor quoted below, 19 instructors—nearly half—used the term "struggle" to describe aspects of students' work in an IBL course.

> Freshman that have done well in school struggle with [math] when they get here. Now they're faced with things that are hard, that they don't know how to do instantly, and they panic—because before, always, they've not had to; it's been easy. But it's okay to not know how to do something, right? I mean, that happens to all great mathematicians, and that's okay now; that's where it gets good. You have something to struggle through. So that course at least gives them a taste of that. It doesn't mean you don't know how to do it, it just means you don't *yet* know how to do it. (*laughs*) (F717; emphasis added for clarity)

Instructors also linked deep engagement with the mathematics to positive affective outcomes. Working on—and eventually succeeding in solving—challenging, non-routine problems helped to develop students' confidence. This gave them a sense of accomplishment and made the resulting mathematical insights personally meaningful.

> They really have to struggle with the problem. And they don't really have the option of giving up and saying, 'Oh, I'll come to class and they'll tell me how to do this.' And so it gives them an opportunity to do more than they might think that they are capable of. (A501)

> When a student discovers something, they discovered it. It's not, you taught it to them, you know, or even, they learned it: they discovered it. They came up with it. And I think that's meaningful to them. ...I think that emphasizing the fact they work hard, and that that work is noticeable, is meaningful to them. ...And I make sure to point it out, maybe not so subtly point it out, when they do that. (A513)

In describing how students learned in IBL courses, instructors made about half as many comments about collaboration as they did about deep engagement. Some emphasized students' roles and responsibilities: "You are going to be talking to your peers, the professors will be at the back of the room" (A506). Many instructors saw value for students in "the accountability of an audience" (A512): both explaining and being explained to helped to clarify ideas and refine understanding. "When you're talking to people, you realize you don't even know anything. So I like group work," said one (A516). Another elaborated:

> A student who has just come up with an idea, at the undergraduate level—I don't believe they understand their idea thoroughly. So I see value in having them explain their idea to their peers, and I see great value in their peers attempting to understand something that's being explained to them by someone who can't do it slickly and without error, and where they have to work to process the information the other person has given them. (A517)

One TA spoke enviously of the collaborative classroom work that she saw benefiting undergraduates by developing both confidence and communication skills:

> They get some confidence to go up to the board and express their ideas, without being afraid to be mistaken or not. And this is something that I haven't [experienced], and I would like to have some kind of confidence doing that. If you are going to do math later on, this is some ability that you would like to have, how to express your ideas—not being afraid to be wrong, but trying to make yourself understandable. I'm really impressed about the reaction of the students. (A503)

Finally a few insightful instructors acknowledged—as did a great many students—that when class was enjoyable, learning was enhanced:

> I think it's really important to make the... class enjoyable for its own sake. I mean, they might have a desire to have a good grade in the class, and that's one kind of a motivation. But it's not as good as liking what you're doing. And you don't work at basketball, you play basketball, right? But you work at mathematics, or at least most students feel that way, that it's work. And so anything you can do to make it seem less like work, especially when you're asking them to work really hard (*laughs*), is a good thing. (F707)

8.3 Critical Teaching Decisions in IBL Classrooms

Instructors offered many thoughts on the processes involved in teaching an IBL course—about 40% of all the coded interview data addressed teaching processes, more than double the level of commentary offered on any other topic. These teaching-related observations were coded into about a dozen specific sub-categories, each addressing different aspects of the instructor's role.

In discussing these data, we are careful to distinguish two components of teaching, adopting the terminology of Speer, Smith and Horvath (2010). First, teachers' *instructional activities* are their "organized and regularly practiced routines for bringing together students and instructional materials to support student learning" (p. 101). These are the activities that teachers plan and direct for their students: lecture, small group problem-solving, whole-class discussion, computer simulations, and so on. Chapter 2 focuses on the choice and conduct of such activities.

Teacher practice, however, involves the "choices and acts" of teaching—the "anticipations..., adjustments..., thinking and decision-making" (p. 101) that instructors do as they implement their chosen activities. Teacher practice thus includes instructors' planning, moment-to-moment classroom decisions, construction and scoring of assessments, and evaluation of students' written and oral work. Two instructors who choose the same activities may thus have very different practices: as Speer and coauthors point out, "Teachers' instructional activities frame and shape, but do not determine their teaching practice" (p. 100). It is such practices that we consider here.

8.3.1 Certain teaching decisions seem to be critical for the optimal success of an IBL course.

At first, the teaching-related data were somewhat puzzling: interviewees often told us in copious detail about specific practices, such as how they graded homework, how they assigned class participation points, or how they selected students to present at the board. It was clear that these details were significant to instructors, but less clear what general lessons we could derive from the particulars of each individual's classroom work.

After some contemplation, we came to recognize that, collectively, these reports outlined a shared set of teaching concerns—a common set of teaching challenges. Individually, each instructor described an idiosyncratic set of practices developed for their own course, student audience, and personal style. These practices were seldom couched in philosophical or pedagogical terms; instead instructors were describing a solution to a specific classroom problem that they never explicitly identified. But as analysts, we began to see that we could abstract from these *individual* narratives the teaching dilemmas that *every* instructor had to resolve. Collectively, these data on teaching processes thus offer great insight into the challenges of IBL teaching, and document a range of classroom-tested strategies that may offer ideas to others.

The issues thus extracted do not differ greatly from those encountered by any teacher: they include choices about curriculum, classroom management, and student assessment. But for IBL classes, some of these become especially salient, for a variety of reasons:

- Most decisions became *more explicit*: Traditional teaching may be based on received wisdom of both students and instructors, often unquestioned, about how things are or should be. In an IBL classroom, student thinking was very much on display: "If the kids are struggling, you are aware of it for sure in IBL" (A506). Instructors often contrasted their IBL experiences with prior teaching experiences: “It makes me think a lot about how little students learn during lecture—makes me think a lot about it” (A515).

- Decision-making became *more dynamic*: With class activities and pacing often in students’ hands, IBL instructors must respond in the moment rather than follow prepared notes. This required alertness to the classroom atmosphere and attention to student responses. As one instructor noted, “The thing that I didn’t anticipate the most was this aspect of how much you really have to be thinking on your feet” (A509).
- Some aspects of class were *more sensitive* to teaching decisions: Because IBL classrooms relied on collaborative learning, decisions that affect student participation or classroom atmosphere might have greater consequences, as this instructor described.

> It’s unlike other courses—you could be a very good teacher in the traditional method and completely ruin this, by saying too much. I think it does work, but it’s more unstable, somehow, if a teacher says too much or is too insulting. A teacher who is too critical of the proofs is real quick going to shut down things. (A507)

- Some issues required that *new solutions* be found: As learning goals shifted away from covering content and toward developing skills in constructing and communicating mathematical arguments, past assessment strategies such as timed, individual tests might no longer measure what instructors valued nor mesh well with the overall course philosophy.

TAs in particular noticed these dynamic and sensitive aspects of IBL instruction, where results subtly depended “on how you tinker with it. You can get it to work well or not work very well at all” (A512).

> I sort of compare [IBL] to a weapon. In the hands of someone good, it’s really good. In the hands of someone who just, you know, expects it to be a magic pill or something, it’s not going to do anything. It’s going to be totally ineffective. (A513)

Often closer in age to the students and not long removed from their own undergraduate experiences, TAs had some empathy with undergraduates, as well as keen knowledge of student thinking from office hours and grading homework. Because they were generally not in charge of class, TAs had more opportunity to observe how students responded to particular instructor moves. TAs who had worked with several different lead instructors could compare the plusses and minuses of these moves, as this TA discussed:

> What [that] instructor did that was so successful is she put the students on stage. She fully accepted her role as moderator—not, you know, [stand] and repeat. …If you ask any of those [other] instructors, they would say that’s what they’re doing. But if you put a camera in the room with everybody [for an] hour, and then they saw that footage, they might see otherwise, you know. ...I’m sitting in the back of the room watching everything happen, so I have a perspective that they don’t have. (A518)

In future work, we will identify and categorize these critical teaching decisions. In the following section, we give some examples, although a full analysis is beyond the scope of this report.

8.3.2 Critical teaching decisions included decisions in planning a course, in adjusting or refining it, and in responding in the moment during a class session.

So far, we have identified several types of critical teaching decisions—some that were made before a course begins (although they might be adjusted or revisited during the term) and others made in the everyday conduct of a course. While preparing a course, instructors generally made decisions about its goals, philosophy and structure:

- Setting curriculum (choice and sequence of problems)
- Setting workload and pace (assignments and expectations)
- Evaluating student learning (setting assessments and grading schemes).

Some of these decisions took on more significance in IBL courses—for example, choices about the nature, frequency and grading of homework assignments. Rather than functioning mainly as a way to assign final grades, checking homework took on a greater formative assessment role, providing students with feedback on their learning and advice on how to improve their work:

> Obviously if they're turning in a lot of problems, at least some of them have to get graded. And I'd like to think that everything gets at least looked at and some feedback gets to them. ….We like to do that very fast. So I spend most of my time actually grading, and trying to give constructive feedback that doesn't give the answers—and that's sometimes hard. (A511)

Other critical decisions addressed the conduct of class time and interactions with students on a day-to-day basis:

- Setting tone and expectations
- Balancing accountability and participation
- Maintaining a positive atmosphere
- Managing interpersonal dynamics
- Providing help
- Evaluating student learning (judging and responding to student work).

Some of these decisions could be anticipated in a general way, but were repeatedly and individually made each time the instructor encountered a problem situation: what to do when a student refused to work at the board, or when the class was stuck on an unproductive idea. An instructor's decision in the moment might depend on particulars such as which student, what mathematical content, and how many minutes of class time remained. Other decisions were sometimes novel and specific to IBL courses, but could be more generally settled. For example, one faculty member described the challenge of using class time productively while waiting for students to write a solution on the board. This challenge was not one he had encountered in previous lecture courses where he did all the writing. Fortunately, a colleague had offered helpful advice, which he deployed right away.

> One thing Calvin told me when we first started is that you're going to be surprised at how long it takes them to get something up on the board and it really does. …While they're getting something up on the board, we'll talk about what they thought about the assignment—and then I'll just see if they have any questions and stuff, just to try to not have [dead time]— because otherwise they just start chatting. (F704)

In professional development workshops and conferences on IBL, we observe that critical teaching decisions of the "in the moment" type seem to be prominent topics. They seem to cause more anxiety among instructors, perhaps because these challenges are not solved in the abstract with a "one size fits all" solution. Conversely, however, many instructors have lower awareness than may be useful of the importance of certain structural decisions, especially those related to assessing students' learning and giving feedback.

8.3.3 Critical teaching decisions were often interlinked, which increased their importance in optimizing an IBL course.

To illustrate how critical teaching decisions were often interlinked, we consider one particular challenge: balancing student participation in class with individual accountability for the material. This topic, raised in 67 observations from 29 different interviews, was important given the collaborative nature of IBL courses.

> I found this in IBL classes and in non-IBL classes, you have students who are clearly falling behind. They stop doing the work or they stop turning in the work, at least…. And you need to motivate them.... I think it's a general phenomenon. But with IBL, it's a little bit more serious if people don't participate, because the majority of the purpose of the class is having the students lead and help each other and sort of struggle through whatever it is they're trying to do that day. And if a significant portion of the class isn't showing up or isn't talking, or hasn't thought about it or done the work, then you have a small group that progresses and then this other group falls even farther behind. (A516)

Instructors' responses to this problem balanced two main principles. They saw value for students' learning in presenting their ideas aloud, so they wanted all students to present. At the same time, however, they recognized differences in students' comfort or confidence in presenting, and they knew it was challenging for some. Thus their strategies were often tailored to particular students, based on their assessment of how much to push and when to back off.

> I almost threw up the first time I had to teach, you know. And so I understand—like, it's very nerve-wracking for certain people to be in front of a group of people. I understand that, but I think that it is a valuable skill and something that you have to get past. ...If somebody was sort of borderline I thought, 'Okay, this guy hasn't presented in a while.' I can tell he's not confident with the material, but he needs to, you know, jump in, and try

> to be an active part of the class—try to think of it as a community of people that are trying to learn together. (A516)

Fairness to other students was a third principle of generally lesser importance. One professor noted, "There were certainly a couple of students who were not pulling their own weight… and I got the sense that they were kind of being dragged along by the core group" (F713). But in general concerns for fairness did not govern instructor choices.

Failure to reach the right balance of participation and accountability had serious consequences. If accountability was too low, participation was throttled, as students had not done enough work to contribute ideas and move the discussion along:

> Not enough people were coming to class prepared for presentation, and because nothing was due on a daily basis, no one was held accountable for not being ready. The class was just stymied. (A517)

However, if accountability was too high, the stakes of presenting were raised too high. The classroom atmosphere became tense and students were intimidated to take a risk or float a tentative idea:

> I think Marjorie last semester counted, like really graded the presentations. She had people in groups and she would give them an A or a B on the presentation—which I think might be good in higher-level classes, but I think her students felt very stressed out by this. So my approach has been more, 'I want you to get to the board and practice, I don't care if you make a mistake…. If you make a mistake, other people have too, and it's good to discuss.' And there are still some kids who don't want to participate—the really shy ones—but most of them, I think, feel pretty comfortable. (F701)

Accountability and participation decisions had consequence for other aspects of class, including:

- The pace at which material was covered: "You have to strike a balance between helping the slow ones along and letting the fast ones [run]. But I feel like we hear from at least half the class on a regular basis. And as long as we structure the classes correctly, we can figure it out so we hear from everyone every couple of weeks." (A511)
- The availability of help outside class, in how students' needs for extra help were met.
- Student engagement: "I think one of the hardest things is accommodating to all levels of students. …Because sometimes I'll see we're spending a really long time on one problem, I'll see people start to get bored, you know, and then they're not participating. With the top students, you have to really make sure that they have something to keep them interested." (A515)

Finally, we offer a sampling of specific strategies that instructors used to adjust the balance between individual accountability and class participation, including strategies for:

- Calling on students in a random, rotated, or pre-selected order, e.g. "I actually programmed my calculator to give me a name." (A509)
- Grading participation, e.g. setting the weight of participation, keeping a record of contributions, informing students of their status to encourage more participation
- Grading homework and setting due dates, e.g. "[I assign] a certain number of problems that they have to turn in, but we only grade the ones that we get to. They know sort of an upper limit of how much to do, but they don't know a lower limit." (A517)
- Communication and coaching, e.g. exhorting the class, encouraging individual students
- Backup plans for days when students were unprepared, e.g. assigning small group work, using a scribe for class-generated ideas: "If people said 'I'll try,' then you'd say 'Okay, get up there and we'll talk you through it.'" (F717)

We develop this example in some detail to illustrate the nature of future work, in which we will flesh out this typology of critical teaching decisions within IBL instruction. We also wish to collect the many specific solutions to each critical teaching decision as a catalog of "go-to" strategies for practitioners to draw upon.

8.4 Benefits and Costs of IBL Teaching for Instructors

In addition to discussing outcomes of IBL instruction for their students, instructors were asked about the outcomes for themselves, both personally and professionally. They reported a variety of benefits and costs from IBL teaching, outlined in the sections below.

8.4.1 Instructors reported numerous personal and professional benefits of IBL teaching.

Instructors reported three main types of benefits from IBL teaching, as itemized in Table 8.2. The most significant was enhancement of their skills, understandings, and values around teaching and learning. Most commonly, instructors described new knowledge of how students understood mathematical ideas. Some gained insight into what was hard about specific material: "In that course I saw exactly what difficulties the students ran into, and they were never what I expected them to be" (F703). Others valued knowing particular students' thinking: "I got to know these students and their ability to think extremely well" (F710).

Observing student learning in an IBL classroom also reinforced instructors' beliefs in the effectiveness of student-centered approaches. Generally these were not radical shifts in their values; most were teaching IBL courses because they thought it ought to work. But personal experience solidified their beliefs and made them feel confident about "handing over the chalk."

> I have a lot more faith in it... I wouldn't have told you that a freshman in honors calculus could have even understood the statement of one of those things. And to see them understanding them and proving them themselves from scratch, it tells me that what I believe students can do is a factor [in] what we teach them. I'm not saying that every student could do that... but the fact that there's that potential there. (A507)

> I'm really learning how they learn—and now I'm realizing that I was really formula- and algorithm-based in teaching and learning.... I mean, there are these different ways to teach these students, so that they can understand the material more conceptually.... I never knew that before. This experience... kind of opened my eyes. (A510)

Table 8.2: Professional and Personal Benefits and Costs of IBL Teaching, Reported by Instructors (43 interviewees)

	Positive observations	Negative observations	Mixed or neutral observations
Enhanced teaching: deepened understanding of students & learning; stronger belief in merits of student-centered learning; nuanced pedagogical skills, patience	100	16	1
Affective benefits: interest, enjoyment, relationships with students, pride in their progress, rewards of meaningful work Affective costs: frustration	43	7	0
Intellectual benefits: stimulation, new insights	15	0	0
Time & effort: more time, effort, mental or emotional work than traditional teaching	0	32	12
Miscellaneous benefits	3	0	7
Total	161	55	20
Percentage of all reported costs and benefits	68%	23%	8%

The third main area of teaching enhancement was in developing more nuanced skills as a teacher: a bigger "tool kit," an "expanded spectrum of possibilities," more "ways to make it work." Specific skills included asking questions in ways that did not give away answers, using wait time to elicit student responses, diagnosing student misconceptions, guiding discussions, pacing a course, structuring students' write-ups of mathematical work, designing problems at the right level of challenge, and guiding students to react to each others' work. A few described quite profound and lasting changes to their teaching style, like this thirty-year classroom veteran:

> There are times when I feel I could never go back to me being at the board with all of the answers. And any time I teach again, it would have to be more, a certain amount of student involvement. I mean, something has changed in the way I am. (F710)

Instructors' negative observations about their growth as teachers (Table 8.2) addressed the lack of opportunity to develop as lecturers (a concern expressed by four TAs at one institution) and some wistfulness about their classroom role. Instructors could see that students benefited from IBL approaches, but some still felt a sense of loss at not being the center of attention: "I like lecturing, I love to talk. But I've found that I've gotten the best results when I just shut up and let students work" (A513).

Affective benefits to instructors included personal enjoyment of the class. Some found it interesting to see what mathematical and pedagogical issues arose each day. Many took pleasure in getting to know students and see them progress, like this TA: "If you are somebody who enjoys watching students learn, you get a much better view of that, a deeper view of that. And you can feel a lot more pride in it. I think it made me a better teacher for any course" (A507).

Others noted a sense of reward from doing meaningful work, like this instructor who had described his interactions with "Julia," a student preparing to be a high school math teacher. "If you can get people like Julia as teachers seeing the joy of reasoning something out, you know, it's a lot more worthwhile than whatever research I do in my office next door and whatever paper I publish that five people read" (F718).

Affective costs to instructors centered primarily on their frustration when students did not prepare for class or were not making progress. This dialogue between two TAs illustrates how frustration was tied to their greater awareness of student thinking.

> TA 1: I mean, I felt like I've wasted an hour and a half of my life, because not only no one came to class prepared, and so we couldn't get anywhere, but then, you know, the way they went and tried to do it in class was not particularly successful for whatever reason. I don't know... sometimes that's really frustrating. But now this year I'm teaching it on my own, and sometimes that's really frustrating for other reasons, right? I mean, yeah, that's part of it—comes with the territory.
>
> TA 2: It's sort of easier to block—to, like, not notice the frustration when you're lecturing, though. I guess that's the…
>
> TA 1: That's true!
>
> TA 2: It's like, in an IBL class you have to deal with it. (*laughs*)
>
> TA 1: And I really do like that—I mean, yeah, if the kids are struggling, you are aware of it for sure in IBL. I think that's great. But at the same time, yeah, there can be days when you're just not making any forward progress. (A506)

They agreed that the benefit of knowing more about how students were thinking came with a built-in cost, as knowing more could sometimes be dismaying or frustrating.

Finally, instructors described it as intellectually rewarding to invent problem sequences for an IBL course, and to think about the mathematics when students introduced an unexpected twist on an idea. "There's always stuff to be thinking about, when the students are stuck. ...Like I feel like there's just a lot more room for strategizing about, you know, how little gas I think that will be just enough to get them to figure it out" (A505).

In sum, instructors reported that IBL teaching offered many benefits, and a small number of costs, in the areas of enhanced teaching, affective or emotional rewards, and intellectual rewards.

8.4.2 The main cost of IBL teaching was greater time and effort.

Instructors made 44 observations about the time and effort involved in teaching IBL courses, most of which described IBL as more work than a traditional course. Preparing problem sequences was a big job, often cited as a deciding factor in whether they would develop a new IBL course. TAs and instructors without TA support pointed out the extra effort they spent grading, to give students rapid and useful feedback on their work, and in office hours, of which IBL students made extensive use. Other instructors noted the demands of an IBL classroom:

> I'm completely focused. ...I'm really trying to understand what they are doing, because I have to correct them.... I have to adjust myself to them, and it's not that they have to adjust themselves to me. So this takes a lot of energy. The class itself is a demanding thing. (F711)

A few instructors felt a strong sense of responsibility for their class, which led them to spend more time thinking about students' learning outside class than they did otherwise, like this TA:

> In [another] class, if I give a good lecture and then I go home, I've done my job for the day. Somehow, here, for my class to be a success is not only determined by me, it's also determined by students. So even if I've done everything I should, and the students just didn't get it, then my job is not done. And I should go home and figure out what I need to do. (A507)

Some of this work was inherent to IBL approaches. For example, the time to grade homework and interact with students in office hours did not lessen over time. But some experienced instructors noted that IBL courses did not require more work once they had established a curriculum and gained some comfort with the approach: "I've tweaked it a lot from year to year, but, really, the hard work is doing it that first time" (F707). Extra time interacting with students and homework was balanced by reduced time preparing for class, they said. Their intellectual work shifted from a focus on the mathematics itself to a focus on how to approach the material and how to interpret students' difficulties with it. The perceived level of effort clearly also depended on instructors' general teaching experience and comfort with the course content. Overall, instructors largely agreed that IBL teaching was somewhat more effort, but also cited the positive outcomes for themselves and their students as making this effort worthwhile.

8.4.3 Early-career instructors reported positive impacts of IBL teaching experience on their career preparation and prospects.

Interviewees cited several career-related outcomes from their IBL teaching experience. Five described definite influences of IBL experience on their current career paths. Three of these described shifts in their career direction toward a greater focus on teaching, and two thought IBL teaching had helped them secure a specific job.

Thirteen interviewees cited IBL experience as providing professional development for future teaching roles. In IBL courses, TAs' role was often less peripheral than in lecture courses, and these leadership roles provided experience in planning and making decisions about a course, "bridging the divide" between being a TA and being a professor, as one put it (A511). "I think it'll make me a better teacher down the road in ways I can't quite anticipate now," said another (A508). A campus leader who mentored graduate student TAs noted, "It stretches them in a way professionally that I think is unexpected for them. ...So to bring all of that with them to wherever they're going to go as a professional educator [and] mathematician—I really like the idea of us peppering the community with that" (F712).

Senior TAs and postdocs who had gone on the job market felt that their IBL teaching experience had been an asset; none of the 14 interviewees discussing this issue felt IBL experience was a liability. Some noted that IBL teaching was viewed by search committees "as a marker that I'm going to do well as a teacher" (A513) or "as a tag that this person would bring interesting ideas" (A517). Others cited positive side effects: strong letters from senior faculty experienced with IBL approaches, teaching awards earned for their IBL work, and the quality of thought they put into the teaching statements that they sent with job applications. As one recent graduate said, "It's helped me rethink teaching and become better at articulating my ideas about teaching. I guess the more different things I do, the more my [notions] of teaching have grown" (A518).

While the sample of IBL Center alumni who have graduated (Ph.D. students) or gone on to permanent positions (postdocs) is too small to draw conclusions about patterns in their job placement, anecdotal evidence suggests that alumni have so far done well in landing positions that make use of their teaching skills. In the future, we hope to follow up with these interviewees to ascertain their career status, satisfaction, and relevance of IBL teaching experience.

8.4.4 At least 85% of instructors wanted to teach with IBL methods in the future.

Overall, teaching with IBL methods was a sufficiently positive experience for the instructors we interviewed that they wished to repeat it. Of 43 instructors interviewed, 31 made explicit statements that they would like to teach again using IBL methods. In some cases, their enthusiasm was profound. "I don't think I've ever been so rewarded in teaching," said one faculty member (F703). Another, who self-described as initially skeptical of the method, said:

> I am not crazy about teaching. I am new at it; I like it to a certain extent— if I am not teaching too much, I usually like it a lot. But this is really—I must say that, in retrospect, this was the best teaching experience I've tried yet. (F711)

Some instructors wanted to teach another course very similar to the one they had recently completed, or one using similar methods for a different mathematical topic. Others described how they would broadly apply "IBL principles" in their other courses—or were already doing so:

> I think it's changed my attitude, in that I don't think I can just lecture the way I did when I taught pre-calculus. I need to get them talking to each other, talking to me more, for them to understand. (A516)

TAs often gave specific examples of activities that they were applying to teaching assignments in lower-level courses but that were derived from their IBL teaching experiences:

- Having students discuss problems in small groups and present problems at the board
- Developing guided inquiry exercises
- Engaging in dialogue with students more often
- Assessing student learning using portfolios or presentations during office hours.

Again, TAs described real and permanent changes to their teaching style. Said one TA, "I already run my office hours differently, just in terms of feeding [students] questions, not jumping in as quickly as I used to—stuff like that" (A611).

Only one person, a faculty member who had recently completed a first IBL teaching experience, reported that he would not teach again using IBL. Two TAs were unsure if they would teach with IBL again; both were early in their teaching careers and wanted a broader range of teaching experiences for comparison before making up their minds about the IBL approach.

Nine instructors made no explicit statement about their interest in teaching again with IBL methods. Of these, four were very experienced senior faculty who had used IBL methods for years, and their intention to continue doing so was not raised in the interview. A fifth was moving on to a non-teaching position, so had no opportunity. A sixth was a mathematics education specialist who expressed interest in other types of student-centered learning. For only three of 43 interviewees is their probable use of IBL in the future truly undetermined.

Thus, counting conservatively, at least 85% of the instructors interviewed expected to use IBL methods again in the future. Of these, some three quarters had recently taken up IBL methods, as faculty new to IBL or as TAs not yet in permanent positions. This is not, of course, a random sample—both instructors and TAs had been selected to participate in their IBL Center's work based on interest in IBL or apparent aptitude for teaching. But the evidence makes clear that an IBL teaching experience was concretely powerful in shaping instructors' later interest in IBL teaching. Indeed, the most important legacy of the IBL Centers may be this cadre of young instructors who have developed strong beliefs in student-centered learning and a nuanced portfolio of teaching skills, who will go on to teach at institutions across the US and world.

8.5 Departmental and Institutional Contexts for IBL Teaching

The mathematics departments that are home to the IBL Centers are distinct in history, mission, and culture, as well as in the size, academic qualifications, and diversity of the undergraduate student body they serve. Moreover, each IBL Center made different choices about how to deploy its grant resources from EAF: at what curricular level, with what groups of students, and for which activities. All these factors result in interesting variations in the number and nature of the IBL courses and in the context in which they operate.

The most noticeable example, evident in the classroom data (see, e.g., Figures 2.1 and 2.2) and clearly influencing student responses, is the predominance within departments of particular IBL teaching "styles." For example, instructors at two institutions used substantial amounts of formal group work in class, while instructors at the other two tended to focus on student presentations at the board. It is clear that these patterns occurred and persisted in part because of how instructors learned about IBL: what colleagues said to them about IBL, what they observed in colleagues' classrooms, and how they adopted or adapted prior versions of the same course as they worked up a course new to them. Relatively few had independent exposure to student-centered instructional approaches via formal professional development, reading, or a collegial community outside the IBL program—thus the IBL approaches in use at a specific campus tended to be somewhat homogeneous. In addition to these style variations, we also noted interesting local variations in

- how faculty and TAs were recruited to teach IBL courses
- how courses were handed off to new instructors
- how instructors new to IBL methods were supported in teaching an IBL course
- whether and how TAs were involved in IBL teaching and how they interacted with faculty
- the nature and extent of collegial community that has grown up around the IBL program
- the nature and extent of communication about IBL with mathematics colleagues not involved in the program.

These processes by which innovative teaching methods are described, transmitted and sustained within a department are very important. They are central to the health and sustainability of the program, its perception in the department and the wider discipline, and the spread of such methods to other instructors and institutions. About 10% of the observations catalogued in the instructor codebook directly address departmental and institutional contexts and processes. Augmenting these data with our own observations at the four Centers, we will offer future analyses of how these contextual factors may support, sustain or hinder IBL teaching.

8.6 Conclusion, Strengths and Limitations of the Instructor Interview Study

Instructors' observations of students provide good corroboration of the student learning outcomes that students reported, from an understandably distinct perspective. Instructors could interpret some of students' gains in light of the skills and understandings students would need in pursuing higher mathematics—gains that students did not yet know would be valuable later on. Instructors proposed interesting differences in who benefited most from IBL methods, some of which were borne out by other studies. They described—again with differing emphasis—the same learning processes that students saw as important in IBL classes, the "twin pillars" of deep engagement and collaboration. In describing their teaching practice, instructors offered many practical strategies that collectively highlight critical teaching decisions that may buttress or weaken these twin pillars, thereby enhancing or limiting the effectiveness with which IBL methods were implemented.

Instructors reported a variety of professional and personal benefits to themselves from teaching IBL courses. Early-career instructors saw their IBL experiences as quite beneficial to their job-hunting process and career prospects. The main costs of IBL teaching were extra time and effort.

In Section 8.5, we noted some differences in departmental and institutional contexts that merit further analysis. But certain similarities in these contexts also inform instructors' perspectives on the issues in this chapter. All the instructors in this study worked in research universities with "very high" research activity (Carnegie Foundation, 2010), and in mathematics Ph.D. programs ranked in the top 50 (Ostriker, Kuh & Voytuk, 2010). They taught well-prepared undergraduates in selective four-year programs, and their IBL program was grant-supported. These factors offered certain affordances as well as some limitations that are not shared by instructors everywhere. As is the case for any interview-based findings, our findings cannot be generalized to all IBL instructors.

Notwithstanding these cautions, the interview sample was large and balanced across a range of career stages and experience levels, and distinct patterns in instructors' observations emerged. Compared with other studies, the instructor experiences reported here are typical of those encountered when instructors take up student-centered teaching methods. Instructors in other settings may place different priority on the outcomes to their students or themselves, or draw different conclusions about (for example) the relative worth of IBL learning experiences versus the workload involved. But the general issues encountered, and the types of observations made, are typical and can usefully guide future studies of instructors in other settings.

8.7 References Cited

Carnegie Foundation for the Advancement of Teaching (2010). *The Carnegie Classification of Institutions of Higher Education.* Stanford, CA. Retrieved March 29, 2011, from http://www.carnegiefoundation.org/classifications/

Laursen, S., Hunter, A.-B., Seymour, E., Thiry, H., & Melton, G. (2010). *Undergraduate research in the sciences: Engaging students in real science.* San Francisco: Jossey Bass.

Ostriker, J. P., Kuh, C. V., & Voytuk, J. A. (eds.) (2010). *A data-based assessment of research-doctorate programs in the United States.* Committee to Assess Research-Doctorate Programs; National Research Council. Washington DC: National Academies Press.

Speer, N., Smith, J. P., III, & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *Journal of Mathematical Behavior*, 29(2), 99-114.

Chapter 9: Summary of Findings

9.1 Introduction

In this chapter, we recap findings from a comprehensive, mixed-methods study of inquiry-based learning (IBL) in college mathematics as implemented at four university IBL Mathematics Centers. Six sub-studies were collectively designed to examine the following questions:

- What are the student outcomes of IBL mathematics courses?
- How do these outcomes vary among student groups, and how do they compare with other types of courses?
- How do these outcomes come about?
- What are the costs and benefits of IBL teaching for instructors and departments?

Chapters 2-8 document findings from each of the sub-studies separately. Viewing results separately is not necessarily the most interesting way to think about them—indeed, they were deliberately designed to provide different lenses on the same problems. But this is a necessary first step in analyzing a massive body of data, that has already unearthed several rich veins of connections to pursue. Here we summarize key findings from all the sub-studies, emphasizing corroborating evidence from multiple lines of inquiry. Wherever possible, we attempt to tell a story that is consistent with all the evidence. We reference specific findings to the chapter or section where that finding is first discussed.

In writing the findings chapters, we wanted to document not just "what we know" but to provide readers with a feel for "how we know what we know." This chapter does not review the evidence and our reasons to trust or doubt it; it only summarizes the claims made. We thus encourage readers to read the original chapters. Nor do we review sample sizes and study methods. Readers may consult the overview of the study design in Chapter 1 (1.5), each chapter for major parameters, and the Appendices[1] for full details. In round numbers, however, the data gathered include 300 hours of classroom observation, 1100 surveys, 220 tests, 3200 student transcripts, and 110 interviews with students, faculty, and graduate teaching assistants (TAs).

9.2 Setting for the Study

The study sites were four university mathematics departments that hosted IBL Mathematics Centers supported by the Educational Advancement Foundation. All four are at research universities (three public, one private) with selective undergraduate enrollment and highly ranked graduate programs in mathematics. They are geographically dispersed and vary in size. The IBL Mathematics Centers were established in 2004, before this study was commissioned. Therefore many of the study parameters were pre-determined, and the study was an observational, not experimental, design, reflecting realistically variable implementations of inquiry-based learning and not an idealized laboratory situation. That circumstance imposes

[1] http://www.colorado.edu/eer/research/steminquiry.html#Reports

some constraints on what we were able to investigate and learn, but it also means that the study findings are relevant to other real-world reform efforts in STEM higher education, offering useful lessons on what can be achieved and by what means.

9.2.1 Courses Studied

Starting in 2004, each Center independently chose certain courses in which to develop and teach IBL mathematics curricula. These courses were well established by the time our research team began the pilot study in 2007 and the main study in 2008. In most cases multiple instructors had been involved in addition to the course originators. Varying levels of professional development and support were available to instructors new to IBL. Some team-taught with an experienced partner, and some attended a summer workshop. Most commonly, however, they visited colleagues' classes before teaching their own, drew upon course materials from their predecessors or colleagues elsewhere, and exchanged ideas informally or at occasional IBL lunches and seminars. TAs often consulted each other and discussed their ideas and experiences.

The courses included in the study addressed a range of student audiences and mathematical topics. Courses for pre-service K-12 teachers focused on deep understanding of the mathematical concepts needed to teach elementary, middle, or secondary schoolchildren. Courses such as cryptology and calculus sought to provide talented first-year students with a stimulating mathematics experience that would draw them into the major. More advanced courses in number theory, analysis, geometry, and discrete math served "mainstream" mathematics majors as well as science and engineering majors. The placement of topics within the mathematics curriculum varied somewhat from campus to campus.

Some IBL courses were offered with parallel non-IBL sections also available. Others—including all of the courses targeted to pre-service K-12 teachers—were offered only in IBL format. Thus the availability of comparison groups for the study was uneven. In the analyses discussed below, data from first-year and advanced courses were combined in describing "math-track" courses, while data from courses for pre-service teachers were analyzed separately.

9.2.2 Inquiry-Based Learning as Implemented at the IBL Centers

We designated course sections as "IBL" or "non-IBL" following the campus Centers' own designations, which were often more pragmatic than philosophical (e.g., IBL courses were those supported by each Center's grant). To establish that IBL designation indeed led to meaningful differences in teaching and learning practices, however, we observed classrooms to directly establish the nature of instruction actually occurring in IBL and non-IBL courses. We found that students in IBL classes spent class time doing and listening to student presentations, working in small groups, participating in whole-class discussion that often arose from a group problem or student-presented solution, and in a few courses, working with peers on computer activities such as modeling and simulations (2.2.1). On average, over 60% of IBL class time was spent on these student-centered activities. In contrast, students in the non-IBL courses spent 87% of their time listening to their instructor talk. While there was substantial variation in practice—

including the occasional "IBL" section that was less than 20% student-centered—overall the grant-supported IBL courses showed very substantial differences in what students experienced in class, compared to lecture-based courses that are traditional in mathematics and other STEM disciplines. That is, classroom practices were clearly distinct between the "reformed" IBL courses and the comparative non-IBL sections.

Drawing on both the classroom observation data (Ch. 2) and student and instructor interviews (7.3; 8.2.6, 8.3), we identify key features of IBL courses in this study by both their statistical frequency and their importance to students and instructors. In IBL courses typical of this project:

- The main work of the course was problem-solving: students solved challenging problems alone or in groups, in and out of class. In class they shared, evaluated and refined their own and each others' solutions.
- The course was driven by a carefully built sequence of problems or proofs, rather than a textbook. The pace of the course was set by students' movement through this sequence, rather than pegged to a pre-set schedule.
- Course goals tended to emphasize development of skills such as problem-solving, communication, and mathematical habits of mind, not just covering specific content.
- Most of class time was spent on student-centered instructional activities. Students or groups of students played a leading role in guiding these activities. Most activities were conducted for just a few minutes at a time: class work tended to change gears often and switched frequently between activities.
- Instructors' main role was not lecturing. They might give mini-lectures to set up or sum up the day's work, introduce a group activity, or provide context for a set of theorems. Instructors (as well as other students) might offer impromptu explanations to respond to a comment or question. They might ask questions to clarify student thinking, refine a proposed solution, give feedback, or elicit such comments from other students.
- Student voices were heard in the classroom: presenting, explaining, arguing, asking questions. Their active participation meant that students themselves had considerable influence on how class time was spent and how fast the class moved through the material.
- This joint responsibility for the depth and progress of the course fostered a collegial atmosphere that placed value on respectful listening and critique and that invited every class member to contribute fruitfully to the mutual development of mathematical ideas. Instructors made efforts to set and maintain this atmosphere.

From instructor[2] and student interviews, we also know something about the behind-the-scenes structure typical of an IBL course. Outside class, much of students' work time—which was substantial—was spent in preparing for class: solving problems or deriving proofs to present or

---

[2] The term instructor refers to faculty and TAs together.

discuss in groups. Because work was due nearly every class, the workload was steady rather than test-driven (3.3.2). Instructors invested substantial time in constructing the "script," or sequence of problems, or in understanding and fine-tuning scripts shared by other instructors. If a script was available, it was rather less work to teach the course than if they developed it themselves. Checking homework took on greater importance for IBL courses, because students' work improved most rapidly when they got timely feedback. Faculty and TAs held many office hours outside class, and students made much use of them, because timely help could be important to making progress, separating "fruitful struggle" from wasted time.

9.3 Student Learning Outcomes

Findings on student learning outcomes were derived from several independent lines of evidence:

- surveys of students' self-assessed learning gains (Ch. 3), which include both numerical ratings and open-ended comments;
- tests given to subsets of math-track and pre-service students (Ch. 5);
- interviews with students (Ch. 7) and their instructors (Ch. 8); and,
- for math-track students, analysis of their grades and course-taking patterns subsequent to an IBL or comparative course (Ch. 6).

On surveys of their self-assessed learning gains, math-track students who took IBL courses reported greater gains than their peers who took non-IBL sections of the same courses (3.2.1). These gains were higher across several areas: cognitive gains, including understanding of mathematical concepts and improved thinking and problem-solving skills; affective gains, including confidence, positive attitude, and persistence; and social gains, including collaboration and comfort in teaching mathematical ideas to others.

For pre-service teachers who took IBL courses, the self-reported learning gains were different (3.2.1). In general, their cognitive and affective gains were lower than those of math-track IBL students. However, their gains in applying mathematical knowledge, collaboration, and comfort in teaching mathematical ideas to others were as strong or stronger than those of math-track IBL students, and clearly higher than those of math-track non-IBL students. These differences likely reflect pre-service teachers' lower general interest in mathematics (3.4.3), but also emphasize their gains in areas that are especially significant for their future work in teaching K-12 students.

Students' write-in comments on surveys reinforce their survey ratings about the breadth and depth of learning they experienced (3.2.2). From both math-track and pre-service courses, twice as many IBL students wrote about learning gains as did non-IBL math-track students. Individual students also wrote much more often of multiple gains from an IBL course. Their voluntary comments heavily emphasized cognitive gains—especially learning more deeply because they had figured things out for themselves. They also described changes in how they learned mathematics and solved problems, including improved learning on their own and from others.

They described affective benefits and new communication skills—less often than they noted cognitive and learning changes, but more often than either type was noted by non-IBL students.

Student interview data further corroborate these findings. Again, students most emphasized cognitive gains (7.2.1), especially the deep and lasting learning that resulted when they worked through ideas for themselves. They saw their gains in thinking and problem-solving skills as transferable to other courses and to life in general. Changes in learning, also commonly reported, were of two types: personal learning changes such as self-awareness, persistence and independence, and greater appreciation for the benefits of collaborative work (7.2.3). Affective gains emphasized confidence, enjoyment and interest (7.2.4). Communication gains included improved writing and speaking about proofs and enhanced abilities to explain ideas and give and receive critique (7.2.5).

Instructor observations of their students' learning aligned well with students' own reports, with understandable differences in perspective (8.2.1-8.2.4). Instructors could better spot gains in communication and understanding the nature of mathematics that reflected students' development as budding mathematicians, but they less easily identified gains in students' personal learning processes (8.2.4).

For pre-service teachers, pre- to post-test score gains on a carefully validated external test of learning mathematics for teaching (LMT) reflected real gains in understanding after an IBL course (5.2.2). In prior work, improved scores on this test have been connected to positive effects on teachers' instruction. The LMT test results thus suggest that IBL courses benefited pre-service teaching students in ways that will benefit their future work as teachers.

With a small sample of math-track students, we gave a "proof test" to compare IBL and non-IBL students' ability to evaluate mathematical arguments and their reasons for judging arguments to be proofs or not. Both groups performed well on the test, and there were no patterns of difference in their overall scores or performance on specific problems (5.3.3). However, there was some evidence that IBL students were more skilled in recognizing valid and invalid arguments (5.3.3) and that they applied more expert-like reasoning in making these evaluations (5.3.6). These results neither contradict nor corroborate the strong cognitive gains reported in other studies, because of limitations in the test samples. The volunteer test-takers were academically strong students—but, as discussed in Section 9.4.2, effects of IBL courses on performance may be most detectable among students whose prior academic record is less strong.

Academic records data from three target courses indicated that IBL students earned grades in subsequent courses that were as good or better than the grades of their non-IBL peers, after controlling for differences in their prior academic records (6.3.1). The pattern favoring IBL students was broadly consistent across different sets of subsequent courses (e.g., next semester, all required, all elective, all IBL), but few of the differences were statistically significant, due to the wide scatter in the grades of both groups. Overall, taking IBL courses may benefit, and certainly does not harm, students' performance in later mathematics courses.

Among students who took mid-level and advanced courses, there were no general differences in pursuit of additional mathematics courses (6.3.2). However, students who took an IBL honors course early in their college career did appear to take more subsequent math courses than did a matched sample of peers from the large lecture-based section. Most of these differences are suggestive rather than definitive because of the small sample size, but the IBL students did complete further IBL courses at a statistically much higher rate. These results imply that IBL honors courses can draw strong students into further study, especially additional IBL courses. Later IBL experiences appeared to neither spur nor deter further study of mathematics.

Overall, several lines of evidence support the claim that students who have a college IBL course grow as mathematicians and as learners in ways that their peers taking non-IBL courses typically do not. The nature and types of the observed cognitive, affective and social gains were very consistent across multiple data sources. Some of IBL students' cognitive gains in reasoning and problem-solving were detectable on tests; and there is some evidence that these gains carry over to benefit students' work in later courses. Of all the learning outcome measures that we compared between IBL and non-IBL students, very few pointed to deficits for IBL students.

9.4 Group Differences in Student Learning Outcomes

Our findings on group differences in student learning outcomes are based on subdividing data sets by student characteristics that may be important. These analyses are based on the same data sets listed in Section 9.3: student surveys, tests, interviews, and academic records. Not every type of group difference could be examined within every data set.

9.4.1 Group Differences by Gender

We examined gender differences in outcomes because, unlike most other STEM fields, women's share of bachelor's degrees in mathematics has declined in the past two decades (NSB, 2010; Figure 2-6), and the decline has been sharpest in research universities (Lutzer et al., 2007).

In survey items on self-assessed learning gains, women in IBL classes reported as high or higher gains than their male classmates across all cognitive, affective and social gains areas (3.2.3). But women in non-IBL classes reported statistically much lower gains than their male classmates in several important domains: understanding concepts, thinking and problem-solving, confidence, and positive attitude toward mathematics. In fact, both men and women reported higher learning gains from IBL courses than from non-IBL courses, but traditional teaching approaches did substantial disservice to women in particular, inhibiting their learning and reducing their confidence. These differences for women were independent of their prior mathematics achievement. Women's spontaneous write-in comments echoed this finding: IBL women wrote over four times as many comments about their cognitive gains, in particular, compared to their non-IBL peers, and also more comments about gains in confidence. IBL approaches appeared to level the playing field for women, compared to traditional lecture-based approaches.

Complementing the data showing strong gains immediately after the course, the academic records data offer perspective on how well these gains may last for women in later courses. IBL women outperformed their non-IBL counterparts on several measures of subsequent grades (6.4.1), as did IBL men (6.4.2). IBL women also took as many or more courses than non-IBL women (6.6.1), though generally fewer than their male IBL classmates (Figures 6.7-6.9). While these patterns were fairly consistent, most differences were not statistically significant.

How women's grades compared to their male peers seemed to depend somewhat on the course level (6.4.3.). Women in the most advanced course held their ground versus their male peers, but women in the first-year course underperformed both their female non-IBL counterparts and their male IBL classmates. Results for women in the mid-level course fell roughly between the other two. First-year IBL women also pursued further IBL courses at lower rates than their male classmates. We suggest that women taking advanced courses are survivors who have learned to thrive independent of the challenges they encounter, while women early in their college careers may be more sensitive to stereotype threat and other differences in real or perceived classroom climate. While the short-term results indicate that IBL experiences particularly benefit women, over the long haul, a single IBL experience may help to close but not erase the gender gap.

We found no gender differences in either the LMT pre/post test score gains (5.2.4) or the proof test results (5.3.4). In fact, few men took the pre-service courses where the LMT was given.

There were no gender differences in the gains reported by the students we interviewed (Table 7.2, 7.4), and very few explicit reports of gender-based differences in their experiences. The fact that gender is such a non-issue in the interview data explains the survey data rather well: from students' perspective, IBL classrooms offered equitable environments where all could succeed and contribute, while non-IBL classrooms were evidently neither effective nor supportive for women. Based on the literature, it is possible that IBL approaches make good use of women's often-greater verbal and collaborative skills, but we have no direct evidence for this in our data.

Overall, IBL experiences appear to be powerful for women, leveling the playing field by eliminating discouraging experiences and negative outcomes of traditionally taught courses.

9.4.2 Group Differences by Prior Achievement

In the pilot study as well as the main study, instructors commonly hypothesized that IBL experiences were most beneficial for students who were not the most academically qualified—students who were good, perhaps, but not great (8.2.5). Thus we examined the data for differences in student outcomes by prior achievement.[3]

On the LMT test given to IBL pre-service teachers, test score gains were anti-correlated with the initial score. That is, students who had the lowest scores on the pre-test improved the most on the post-test. Those with less than 50% correct on the pre-test improved by more points than their initially high-scoring classmates (>75% correct) (5.2.3). Students with medium pre-test scores (51-75%) improved more than high-scoring but less than low-scoring peers.

[3] Different data were used to distinguish prior achievement in each sub-study: for surveys, self-reported overall GPA and expected course grade; for the LMT, pre-test score; for academic records, prior math GPA. We do not know how these relate to instructors' characterizations of "strong" and "weak" students.

This finding matches students' self-report of learning gains, where IBL pre-service teachers with lower overall GPAs reported higher gains than their classmates with higher grades (3.2.6), mirroring the pattern seen on the LMT test.

For math-track students, the pattern among IBL students was less striking, but similar in that the students with the lowest GPAs reported higher gains than the middle GPA group (3.2.6). This was interestingly different from the pattern in non-IBL classes, where students who had the highest GPAs reported the highest gains, and low-GPA students reported the lowest gains.

In the academic records data, we also saw that, among students who entered with low math GPAs (<2.5), IBL students generally outperformed their non-IBL peers in subsequent classes (6.5.1). IBL students who entered with higher math grades also did as well or better than non-IBL peers in later courses (6.5.2, 6.5.3), but the improvement for previously low-achieving IBL students was striking (6.5.4.).

Overall, there was some evidence that non-IBL courses tended to reinforce prior achievement patterns, helping the "rich" to get "richer." In contrast, IBL courses seemed to offer an extra boost to lower-achieving students, especially among the pre-service teachers. Yet there was no evidence of harm done to the strongest students. Indeed, the highest-achieving students may have been encouraged by an IBL experience to take more mathematics courses, especially more IBL courses (6.7.2)—again, consistent with instructor observations that strong students found the IBL approach stimulating (8.2.5). Instructors did have reservations about the benefits of IBL for the "weakest" students, and our analyses may not detect a small number of students who struggled noticeably. Moreover, if these students dropped the course, they were not represented in our post-course data sets.

9.4.3 Group Differences by Experience Level

In the pilot study, first-year students were particularly enthusiastic about how IBL courses had enhanced their learning in other courses. Thus we investigated differences in student outcomes for students taking IBL courses earlier or later in their college careers.

On surveys, first-year math-track students who took IBL courses reported higher gains than did IBL students later in their careers, across several areas: mathematical thinking, persistence, and collaboration (3.2.4). Moreover, both first-year and mid-career (sophomore/junior) students reported higher gains in these areas, as well as confidence and positive attitude about mathematics, compared to their non-IBL peers. Even the late-stage (senior and graduate) students reported higher affective and collaboration gains than their non-IBL peers.

A similar pattern was found when survey data were differentiated by the number of prior college mathematics courses taken (3.2.5). Among IBL students, less experienced students reported higher gains than more experienced colleagues, but this was not the case for non-IBL students. Gains enhanced by IBL experience included cognitive, affective, and social gains for novice math students, but only affective and social gains for students with more math background.

In interviews, first-year math-track students reported more gains overall than advanced math-track and pre-service teaching students (Table 7.3; 7.4). Consistent with the survey results, they particularly emphasized cognitive gains, changes in their understanding of the nature of mathematics, and affective gains including confidence and enjoyment.

In sum, several lines of evidence indicate that IBL experiences were more powerful for students earlier in their college career. This finding is also consistent with data from an IBL first-year honors course on students' later course-taking patterns (Figure 6.9). A first-year IBL experience may contrast strikingly with students' high school work, and changes in students' approaches to learning or studying may influence their work in later courses in mathematics and other fields.

9.4.4 Other Group Differences

When considering other student sub-groups, we found very few other differences, and no systematic patterns of difference. We detected no meaningful differences in the outcomes of IBL courses for students of varying race, ethnicity, or academic major. However, the sub-sample sizes were small for students of color and for non-mathematics majors; these study sites did not provide good tests of these issues. In general, however, we have no evidence that IBL methods did not work equally well for students of different personal and academic backgrounds.

9.5 Student Attitudinal Outcomes

Student surveys explored several attitudinal variables that characterized students' mathematical beliefs, motivation and strategies for learning and problem-solving. Interview data on attitudes, especially in relation to student resistance to IBL approaches, have not yet been analyzed.

9.5.1 Characterization of Students' Beliefs, Motivations, and Learning Strategies

Based on their pre-course survey responses, math-track students in both IBL and non-IBL classes had strong interest in mathematics, and high motivation in both intrinsic (internal) and extrinsic (grades, future plans) dimensions (3.4.2, Table 3.5). They held fairly sophisticated views of mathematical problem-solving as a constructive and logic-based process. While IBL and non-IBL students were alike in many ways, IBL students more often rated mathematics as a personal, not just academic interest. They were also more confident and had greater preference for group work. These differences confirm that there is some preferential selection (by advising and/or self-selection) of certain types of students into IBL courses, as students and faculty also told us.

Compared with math-track students, pre-service teachers were less interested in mathematics, less likely to enjoy it, and more extrinsically motivated (3.4.3). Their beliefs about learning and problem-solving were somewhat more novice-like, viewing learning as more instructor-driven and problem-solving as more about confirming truths and practicing procedures than about discovering ideas. However, they believed more in the value of group work, made more use of it in their own studying, and were more interested in teaching and communicating mathematics.

9.5.2 Attitudinal Changes Following an IBL Course

In general, the changes in these attitudinal variables from pre- to post-course survey were modest for all groups (3.4.4). They were also modestly but positively correlated with student learning gains (3.4.13), showing that attitudinal changes and learning outcomes are related.

For non-IBL math-track students, attitudinal changes were mixed, but more negative than positive (3.4.5). After their course, students reported lower confidence and enjoyment, less willingness to study hard for a math course, and less strong beliefs in rigorous reasoning as a general problem-solving approach. But for IBL students, most of the changes were positive: stronger personal interest in mathematics and in communicating it, stronger beliefs in proving as a constructive and creative activity, and stronger beliefs in and use of collaborative learning.

Among pre-service teachers, attitudinal changes were mixed (3.4.6). Following an IBL course, they placed less emphasis on extrinsic goals and instructor-driven instruction, suggesting some maturation of their approach to learning mathematics. However, they did not gain in confidence, and lost some ground in their use of the self-regulatory learning strategies that are used by successful learners.

For women, there were small positive changes in confidence and motivation following an IBL course, contrasting with larger negative changes in confidence, collaboration, and use of effective learning strategies for women who took traditionally taught courses (3.4.7, 3.4.8). These findings align well with the learning gains observed for women (9.4.1): again, IBL approaches appeared to remedy problems with traditional lecture-based teaching that were particularly detrimental to college women's interest and confidence in mathematics.

Among first-year students, there were enhancements to students' strategies for learning and problem-solving, while these did not change for older students (3.4.9-3.4.11). But among older students, positive changes in interest, motivation and confidence were observed, which were instead modestly negative for the first-year students. We suggest that early IBL experiences have an influence on students' approach to learning that may be powerful if it carries over to other college work. Later IBL experiences may not shift students' well-established study habits and beliefs, but may revive their interest in mathematics, as some interviewees suggested (7.2.4).

Overall, IBL math courses tended to promote more sophisticated and expert-like views of mathematics and more interactive approaches to learning. In contrast, traditional mathematics courses appeared to weaken students' confidence and enjoyment, and did not help them to develop expert-like views or skillful practices for studying college mathematics.

9.6 Teaching and Learning Processes

Several lines of evidence clarify the teaching and learning processes important in IBL courses. Clear differences in IBL and non-IBL student outcomes (9.3) mirrored the clear differences between IBL and non-IBL course practices (9.2.2). More importantly, we can directly link student outcomes to course practices: student learning gains correlated statistically significantly with the fraction of class time spent doing student-centered activities (small group work, student presentation, computer work, and discussion), and anti-correlated with the fraction of time spent listening to instructors talk (4.3.1). Similar correlations were found for the relation of learning outcomes to the proportion of class time that was student- or instructor-led (4.3.1), and for variables that reflect how students and instructors interact and share responsibility for the course (4.4). Moreover, statistical modeling shows that the degree of student-centered class time was the strongest strong predictor of student learning as measured by our broadest learning indicator, survey learning gains. When observation data are not available, the binary IBL/non-IBL label was a good predictor of learning.

Second, students themselves reported in some detail on how particular course practices supported their learning. On surveys, IBL math-track students cite several practices as "helping me learn" to statistically higher degrees than cited by their non-IBL peers: the overall approach; their own active participation; and interactions with the instructor, the TA, and their peers (3.3.1, 3.3.2). Non-IBL students cited tests as important for their learning, while IBL students found other types of assignments more helpful. Both groups cited individual effort as important to their learning. IBL pre-service teachers emphasized a somewhat different mix of experiences, including their own active participation and interaction with instructors, but also tests.

Interactive and collaborative course experiences were especially important for women (3.3.3) and first-year students (3.3.4), which may help to explain the strong learning outcomes for these groups (9.4.1, 9.4.2). There were no clear patterns of difference in the experiences of other student sub-groups, consistent with the lack of clear patterns in their learning outcomes (9.4.4).

Finally, in interviews, student discussion of their learning processes emphasized the twin pillars of deep engagement with mathematical ideas and collaboration with others (7.3.1). Deep engagement fostered deep understanding; it rested on both students' individual effort and the assignment of meaningful problem-solving tasks that were not mere "busy work." Collaboration was integral to IBL courses, whether as structured small group work, whole-class discussion, or out-of-class informal group work. Students found it efficient and useful to tackle hard problems with multiple brains, and they learned from explaining their ideas and trying to understand others. The twin pillars reinforced each other: after struggling with a problem individually, students were well prepared to contribute meaningfully during class, and interested in the solutions that others proposed. Collaboration in turn motivated them to complete the individual work. It also made class enjoyable, encouraged clear thinking, and built communication skills.

Instructors' observations of student learning processes confirm the twin pillars as critical and link them to students' positive cognitive and affective outcomes (8.2.6). Instructors' descriptions of their teaching practices enable us to identify critical teaching decisions that can sensitively influence the success of an IBL course. Their choices about course materials, assessment, classroom dynamics, and other factors may aid, abet, or interfere with the central learning processes of deep engagement and collaboration (8.3).

Overall, surveys and interviews provided strong and consistent evidence about the dual importance of individual engagement and collaborative learning processes in IBL courses. While deep individual engagement was a cornerstone of Moore's original method, the importance of collaborative learning reflects a deliberate shift in modern instructional practices, resulting in enriched student learning, growth in collaboration and communication skills, and an enjoyable experience for students and instructors alike.

9.7 Outcomes for IBL Instructors

Instructors reported numerous professional and personal benefits of teaching with IBL methods, which outnumbered their costs by a 3:1 ratio (8.4.1). Chief among these were enhancements to their teaching: deeper understanding of students and learning; stronger beliefs in the value of student-centered learning; and a larger and more nuanced portfolio of teaching skills. Other benefits included intellectual stimulation, and the affective benefits of interest, enjoyment, and pride in their students' progress. The main costs were time and effort (8.4.2), and occasional frustration with students' lack of progress or participation (8.4.1).

Early-career mathematicians felt that their career preparation and prospects were enhanced by IBL teaching experience (8.4.3). For a few, IBL experiences had prompted definite shifts in their career path; but more common were reports of professional development that was seen to prepare them for future teaching roles. Of fourteen who had gone on the job market, all reported that their IBL background had been an asset rather than a liability: a signal to departments that they were serious about teaching, a source of strong recommendation letters, and a thought-provoking experience that helped them write articulately about teaching in their job applications.

Of the instructors interviewed, at least 85% wanted to teach with IBL methods again (8.4.4). They described profound and permanent changes in their teaching styles and beliefs about what and how students learned. These influences on their teaching approach carried beyond the original courses where they had learned IBL methods, as they took what they called "IBL principles" to other courses, including high-enrollment lower-level courses such as calculus and pre-calculus. Most of their learning came from on-the-job experience, although in most departments a loose-knit IBL community had emerged that could offer support and ideas.

Overall, IBL teaching experiences were rewarding for instructors. As professional development experiences, they permanently shifted instructors' beliefs and practices toward student-centered approaches known to improve student learning (Bransford, Brown & Cocking, 1999; Prince, 2004). The most important legacy of the project may be this cadre of young instructors who gained IBL teaching experience at the Centers and are now moving on to teach at institutions across the United States and in other countries.

9.8 Outcomes for the Project Viewed as a Reform Effort

In the preceding sections, we have summarized evidence about the nature of the teaching methods implemented across the IBL Centers and the quality of the resulting student experience. Considered broadly, the teaching and learning methods implemented at the IBL Mathematics

Centers were consistent with evidence and best practices from research on the learning sciences. Their application in these contexts yielded positive outcomes for many students. Early-career instructors in particular reported professionally developmental experiences that appear likely to shape their future practices, thus amplifying the impact of the Centers' work over time. These outcomes for students and teachers are one way to view the overall impact of the IBL Centers.

From another perspective, however, we can view the Centers' work collectively and examine the magnitude of its impact as a single reform effort. That is, we can ask, How big a difference has this project made? Based on typical course offerings, class sizes, and data about student enrollment in multiple IBL courses, we estimate that about 425 unique, non-repeating students had an IBL experience each year during our study period. Of these, about 215 were math majors and about 180 were pre-service K-8 teachers (the rest were in other fields). Over the same period, these four departments graduated about 500 mathematics majors each year (NCES, 2011). Thus, on the order of 40% of all mathematics majors graduating from the four Centers may have had an IBL experience, although clearly this percentage varied substantially depending on how departments chose to deploy their IBL resources.

Likewise, we can compare the number of affected IBL pre-service teachers with the number of education degrees and/or post-baccalaureate certifications awarded each year—about 160 at the two institutions whose IBL programs targeted pre-service teachers (NCES, 2011). (State-by-state differences in teacher certification procedures mean that not all pre-service teachers are accounted for by the NCES data set.) Considering as well the course requirements for students taking the targeted pre-service courses, it appears that essentially all students preparing for elementary/middle school teaching at these two campuses had an IBL experience.

Another measure of impact is the sustainability of the IBL Centers' effort over time. In a time of economic hardships for higher education, all the Centers had taken measures to protect aspects of their IBL program from budget cuts. Some aspects of the IBL programs were seen to be "here to stay," yet leaders also testified to aspects that were fragile, perhaps unlikely to survive a loss of grant funding. In times of economic retrenchment and failing support for U.S. higher education, the jury is still out as to the sustainability of these reform efforts, with or without grant support.

9.9 Issues for Future Research

We are confident that findings from this study will make real contributions to the literature on student outcomes of IBL for mathematics majors and for pre-service teachers (Hough, 2010a,b). Likewise, we will add to the scanty body of evidence on teaching practices in college mathematics (Speer, Smith & Horvath, 2010) and on factors that shape STEM higher education reform (Fairweather, 2008). We look forward to moving these analyses into peer-reviewed scholarship, and to communicating with practitioners who can make use of these findings. We will be pleased if the tools developed for this study are useful in others' research and practice.

Our data sets are voluminous, and this report does not do full justice to them. Some analyses are well advanced while other data sets are just partially explored. In Chapters 2-8, we identified some issues that merit further analysis of the existing data sets. Chief among these are:

- Exploring linkages between student learning outcomes and classroom practices, including application of hierarchical linear modeling approaches that can account for course-level variation separately from student-level variation
- Elucidating the roles of particular course elements in student learning, such as pace and workload, curriculum, assignments, assessments and grading, access to help, etc.
- Extracting and categorizing lessons on critical teaching decisions that will aid other instructors in refining their own IBL practices
- Analyzing contextual factors that influence the health and sustainability of student-centered teaching practices in mathematics departments.

This study also raises several issues that merit further investigation beyond this study. We recommend that future study of IBL mathematics at the undergraduate level include:

- Well-controlled comparative studies of student learning of mathematical concepts and thinking skills, using tests, oral exams, or other forms of student work. Data comparing self-report to externally validated measures would be useful.
- Development of assessments to measure students' growth in valued mathematical habits of mind such as proving skills, critical thinking and creativity
- Development, dissemination and action research on classroom assessments that offer alternatives to timed, individual tests
- Investigation of student outcomes of IBL as implemented in other settings, especially:
  - In other types of courses, especially lower-level and general education courses
  - In other types of institutions, including comprehensive universities, liberal arts and community colleges
  - With student populations that are diverse in race, ethnicity, and academic background
  - By instructors of different background and expertise.

9.10 Conclusion

The approaches implemented at the IBL Mathematics Centers benefited students in multiple, profound, and perhaps lasting ways. Learning gains and attitudinal changes were especially positive for groups that are often under-served by traditional lecture-based approaches, including women and lower-achieving students. First-year and less mathematically experienced students also benefited particularly. Yet there was no evidence of negative consequences of IBL for men, high-achieving students, older and more experienced students: these groups too made gains greater than their non-IBL peers.

The positive outcomes for students were linked to classroom practices that emphasized deep engagement with mathematical ideas and collaborative exploration of these ideas. IBL classrooms offered equal learning opportunities for men and women and motivated students to invest their own effort to advance class progress. Instructors also benefited from their IBL teaching experiences and made lasting changes to their teaching practice. On the whole, the teaching and learning methods implemented at the IBL Mathematics Centers were broadly consistent with evidence and best practices from research on the learning sciences. Our results augment that body of evidence with support for student-centered approaches to undergraduate mathematics education.

9.11 Acknowledgments

We thank a very long list of individuals (1.7) for their assistance and for the pleasure of working with them. We thank the Educational Advancement Foundation (Austin, TX) for support of this work. All conclusions are the authors' own and do not necessarily reflect the views of the EAF.

9.12 References Cited

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.) (1999). *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: National Academies Press.

Fairweather, J. (2008). Linking evidence and promising practices in science, technology, engineering, and mathematics (STEM) undergraduate education: A status report for the National Academies Research Council Board of Science Education. Retrieved 4/23/2011 from http://www7.nationalacademies.org/bose/Fairweather_CommissionedPaper.pdf

Hough, S. (2010a). The effects of the use of inquiry-based learning in undergraduate mathematics on student outcomes: A review of the literature. [Report prepared for Ethnography & Evaluation Research, University of Colorado at Boulder.]

Hough, S. (2010b). The effects of the use of inquiry-based learning in undergraduate mathematics on prospective teacher outcomes: A review of the literature. [Report prepared for Ethnography & Evaluation Research, University of Colorado at Boulder.]

Lutzer, D. J., Rodi, S. B., Kirkman, E. E., & Maxwell, J. W. (2007). *Statistical abstract of undergraduate programs in the mathematical sciences in the United States: Fall 2005 CBMS Survey*. Washington, DC: American Mathematical Society.

National Science Board (2010). *Science and Engineering Indicators 2010*. Arlington, VA: National Science Foundation (NSB 10-01). Retrieved 1/27/2011 from http://www.nsf.gov/statistics/seind10/start.htm

National Center for Education Statistics (NCES) (2011). Integrated Postsecondary Education Data System (IPEDS). Department of Education. Office of Educational Research and Improvement. Data retrieved 4/25/11 from http://nces.ed.gov/ipeds/datacenter/ .

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*(3), 223-231.

Speer, N., Smith, J. P., III, & Horvath, A. (2010). Collegiate mathematics teaching: An unexamined practice. *Journal of Mathematical Behavior*, 29(2), 99-114.