From innovation to implementation: Multi-institution pedagogical reform in undergraduate mathematics

SANDRA L. LAURSEN

Ethnography & Evaluation Research, University of Colorado Boulder, Boulder, Colorado, USA
Corresponding author: Sandra.Laursen@colorado.edu

Keywords: inquiry-based learning; mathematics; undergraduate; mixed methods

The use of student-centered teaching approaches improves student learning and persistence in undergraduate science, engineering and mathematics, but most prior studies have investigated these reforms on small scales and in well-understood conditions. A study of inquiry-based learning (IBL) as applied in undergraduate mathematics at four U.S. research universities demonstrates that such reforms are effective when applied on a multi-course, multi-institution scale that can make a real impact on student outcomes. Here we highlight three key points relevant to research and practice. First, despite variation in the nature and quality of IBL implementation across some 40 courses and 100 course sections studied, positive student outcomes are detected relative to traditionally taught courses. The use of IBL methods benefits women, first-year and initially lower-achieving students in particular ways. Second, the research design appropriate for such studies will necessarily differ from those possible in single courses, requiring general measures that can be used with multiple courses and audiences. Finally, sizable investment in course-specific reforms does not assure that they will remain in place, but professional development of early-career instructors is a powerful byproduct that spreads the method to new courses and institutions and thus helps to broaden faculty use of research-based reforms.

1 Introduction

Mounting and persuasive evidence shows that the use of student-centered "active learning" approaches to science and mathematics instruction improves students' learning and persistence [1-4]. Yet most undergraduates do not experience these proven, "high-impact" educational practices, and those most likely to benefit—minority, poor, and first-generation university students—are least likely to do so [5]. Today, the bottleneck in improving undergraduate education in the so-called STEM (science, technology, engineering and mathematics) fields is not lack of evidence that student-centered approaches work, but lack of uptake: adoption by many faculty at all types of institutions, and institutional commitment to sustaining these approaches [6-9]. That is, "The problem in STEM education lies less in not knowing what works and more in getting people to use proven techniques" [6, p. 28].

Taking such a view of STEM higher education, I argue, shifts the task of both practitioners and researchers. This is analogous to the process for approval of new drugs or medical devices in many countries: researchers first focus on the safety and efficacy of the intervention, but once this is satisfactorily demonstrated, they turn to large clinical trials to understand how well treatments work in different sub-populations, identify less common or long-term effects, and compare treatment regimes. So too must educators' focus shift from demonstrating proof of concept in known conditions to applying and understanding the outcomes of educational interventions on a large scale, for students of diverse backgrounds and under conditions of real-world variability.

Our team has studied an example of reform on this larger scale of multiple courses and multiple institutions, the application of inquiry-based learning (IBL) to undergraduate mathematics at four U.S. research universities with privately funded "IBL Math Centers." Similar to other inductive teaching approaches [10], IBL approaches in mathematics involve students in working out ill-structured but meaningful challenges. Students construct, analyze and critique mathematical arguments, building toward major concepts through a carefully designed sequence of problems rather than a textbook. Thus students' ideas and explanations define and drive progress through the curriculum. Class time is used not for introducing ideas, but for working them through, as students present and discuss problem solutions individually or during structured small group work, while instructors guide and monitor this process.

In U.S. mathematics, IBL approaches have developed mainly through collegial sharing and adaptation of the Socratic methods of renowned topologist and teacher R. L. Moore (1882-1974). In Moore's courses, students worked entirely independently, given axioms but no lectures or textbook, and presented their proofs in class; a student who had not solved a proof was to leave class rather than listen to the solution. Modern incarnations are importantly modified to incorporate peer interaction, foster a positive learning environment, and welcome a broader talent pool. Today the name IBL describes a set of approaches that emphasize learning by participating in authentic mathematical practice and that incorporate well-structured collaborative work as well as individual work on meaningful mathematical tasks [11]. These are consistent with but not historically grounded in cognitive science or learning theory.

In 2004, the IBL Math Centers were established to promote IBL methods through course development, dissemination, and professional development of graduate students and postdoctoral scholars. The Centers were located at U.S. research universities (three public, one private) with selective undergraduate enrollment and highly ranked graduate programs in mathematics. Faculty have offered over 40 IBL courses across multiple topics (analysis, number theory, discrete math, cryptology, multivariable calculus, differential equations, foundations of school mathematics) to varied audiences (first-year honors students, upper-division math majors, pre-service primary and secondary teachers). Each university independently selected courses and prepared its instructors. The resulting suite of courses shared a general philosophy and drew from a common set of teaching approaches, but were diverse in structure, content, and implementation. While posing challenges for study design, these highly variable conditions are typical of any real-world case of educational reform.

2 Study Methods

Our study examined student outcomes of the IBL courses, their variation among student subgroups, and teaching and learning processes in the courses. Because neither students nor instructors could be assigned to IBL methods, the study used a quasi-experimental design. Given the large range of courses and audiences, as well as variability in how faculty implemented IBL, we applied general measures, rather than relying on course-specific learning assessments, to examine student outcomes. These measures included:

- pre/post surveys of attitudes, beliefs and approaches to learning mathematics
- post-course surveys of learning gains, via the SALG-M, based on the Student Assessment of their Learning Gains (salgsite.org) [12]

- course grades and course completions subsequent to the IBL or comparative course, from academic records
- mathematics content tests applied to small subsets of specific courses
- interviews of students, faculty and graduate teaching assistants about their learning and teaching experiences, outcomes and processes
- classroom observations of the use of class time and the learning environment.

Each of these methods is described in detail in [13]. In all, the data include some 300 hours of classroom observation, 1100 surveys, 220 tests, 3200 student transcripts, and 110 interviews, gathered from over 100 course sections at four campuses over two academic years. Comparative data for "non-IBL" sections of the same course were available only for "math-track" courses, those for students seeking a mathematics degree or a science/engineering degree with high math requirements. Pre-measures were applied to control for differences in IBL and non-IBL student samples. Data from IBL math courses targeted to pre-service schoolteachers were analysed separately due to the differing mathematics backgrounds, goals and attitudes of these students [14].

3 What is IBL? Characterizing the Educational Intervention

To argue that the Centers' work is a multi-institution "reform," we must establish that the instruction was both student-centered and distinct in IBL vs. non-IBL courses. We accepted the Centers' labelling of course sections as IBL or non-IBL, but used classroom observation to describe what actually took place, corroborated by student surveys and interviews and by instructors' descriptions of their practice and rationale.

These data showed that, during IBL classes, students gave and listened to student presentations, worked in small groups, used computers for simulation and modelling, and discussed ideas that arose from these experiences. On average over 60% of IBL class time was spent on such student-centered activities. In contrast, students in non-IBL courses spent 87% of class time listening to their instructors talk. IBL courses also featured greater degrees of student leadership, more variety in classroom activities, and higher frequencies of student question-asking. Observers rated IBL courses higher for a supportive classroom environment, students' intellectual contributions, and in-class feedback to students on their work. Interestingly, ratings of instructor-centered behaviors did not notably differ, suggesting that the main distinction between IBL and non-IBL classrooms lies in instructors' choice of instructional activities rather than in their intent as teachers or interest in student learning.

Among IBL courses, there was substantial variation in the extent and nature of student-centered activity (Figure 1) as well as in instructor skill at executing these methods. Nonetheless, IBL courses commonly featured:

- learning goals focused on problem-solving and communication
- a curriculum driven by a carefully constructed sequence of problems or proofs, driving toward a small number of "big ideas"
- course pace set by students' progress through this sequence
- class time used for a mix of active and collaborative problem-solving tasks
- instructors who guided student work instead of delivering information.

Outside class, interviews reveal that much of students' substantial work time was spent in preparing for class: solving problems or deriving proofs to present or discuss. Because work was due nearly every class, the workload was steady rather than test-driven. Instructors invested their own time in constructing the "script," or problem sequence, or in adapting scripts shared by other instructors. Checking homework also took on more importance for IBL courses, because students' work improved most rapidly when they got timely feedback. Students made much use of their instructors' office hours outside class, because timely help could be important to making progress.

In addition to detecting clear differences in instructional practice between IBL and non-IBL courses, we also saw patterns in preferred campus styles of IBL, and in courses taught for certain audiences. For example, structured small group work was more common in courses for first-year students and for pre-service teachers, while individual student presentations were more common in advanced courses. These patterns reflect instructors' beliefs, shared in interviews, about the balance of support and independence appropriate for students at different developmental levels.

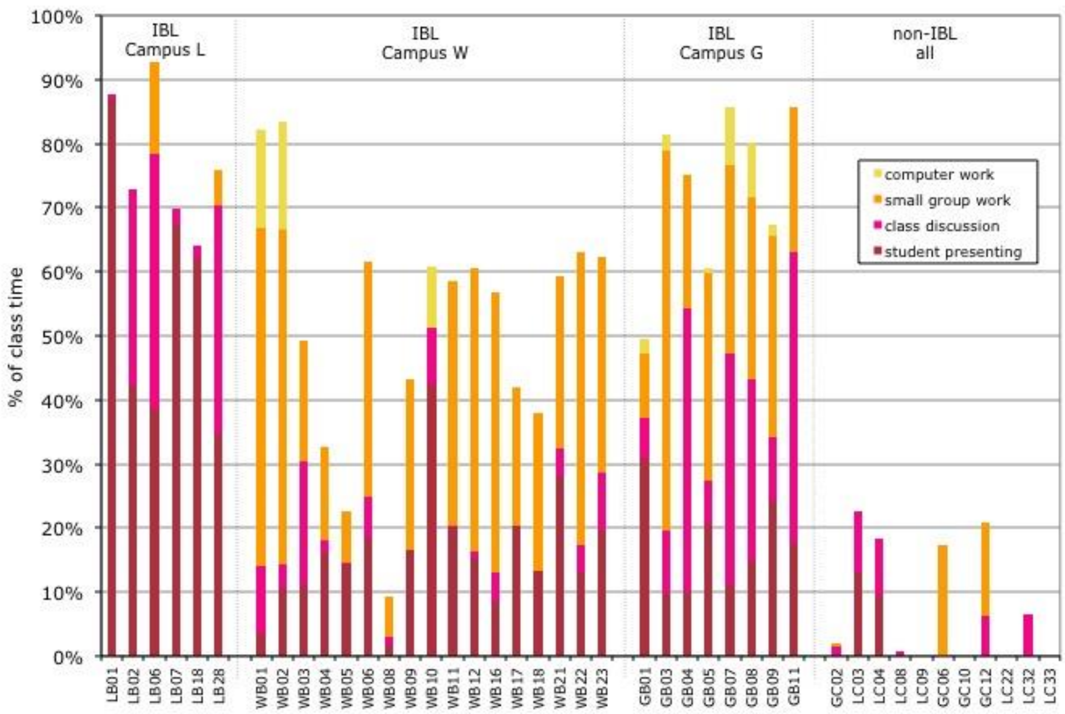

Figure 1: Distribution of classroom learning activities for individual courses, categorized by an external observer, as percentage of all class time observed. Eleven non-IBL and 31 IBL course sections on three campuses were observed for multiple sessions totalling (on average) 313 minutes each; course sections are denoted by codes. Four student-centered categories of learning activity are shown; the remaining time was categorized in two instructor-centered categories or as non-instructional business.

4 Results: Student Outcomes

Knowing that there are real and meaningful differences in what students experience in the classroom, we can then examine student outcomes for differences that may be attributed to IBL instructional approaches. Focusing mainly on "math-track" courses, we present results for each type of measurement in the study. We report differences among IBL and non-IBL students, then consider important sub-groups of students. Further details can be found in our comprehensive report [13].

4.1 Learning Gains

On the SALG-M, IBL math-track students reported greater learning gains than their non-IBL peers, on every measure, including cognitive gains in mathematical thinking and understanding; affective gains in confidence, persistence, and positive attitude about math; and collaborative gains in working with others, seeking help, and appreciating different perspectives (Figure 2). Students' comments corroborate these gains, especially the deep and lasting learning that came from working through ideas for themselves. They saw gains in thinking and problem-solving skills as transferable to other courses and to life in general. Changes in learning included personal learning changes such as self-awareness, persistence and independence, and greater appreciation for the benefits of collaborative work. Instructors' observations concurred, but they could better spot gains that reflected students' growth as budding mathematicians, such as communication skills and understanding the nature of mathematics.

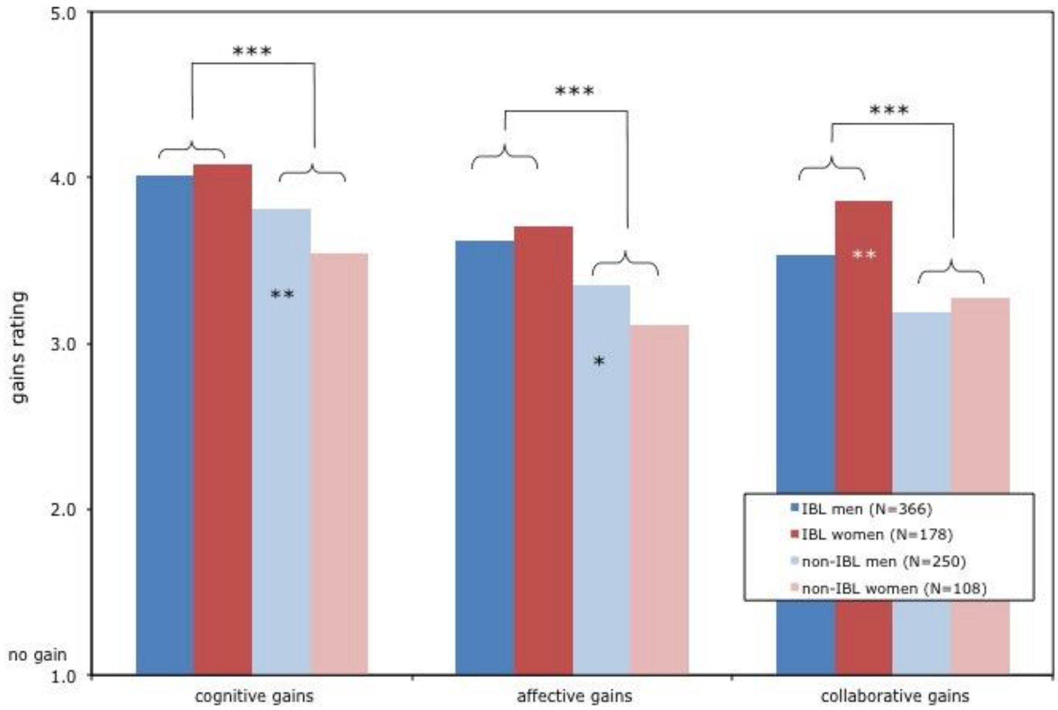

Figure 2: Mean learning gains for men and women students from IBL and non-IBL courses. The mean for cognitive gains represents a composite of 4 items (math concepts, math thinking, application, and teaching), and for affective gains a composite of 3 items (confidence, positive attitude, persistence). Collaborative gains represents 1 item (working with others). Asterisks denote statistical significance: *** $p<0.001$; ** $p<0.01$; * $p<0.05$. In-group comparisons (e.g. men vs. women in IBL courses) are marked on the bars; between-group comparisons (IBL vs. non-IBL) above the bars.

The gains are more dramatic when separated by gender (Figure 2). Women in non-IBL courses reported substantially lower cognitive and affective gains than did their male classmates. But in IBL courses, women's gains were statistically identical to those of men. To more rigorously model this result, we first tested for differences in the IBL and non-IBL groups' initial characteristics using variables such as academic background and prior mathematics experience. Using propensity analysis [15], we created a combined variable which, when used as a covariate, held these characteristics constant in our Analysis of Variance (ANOVA). In the same model, we also used three attitudinal covariates derived from factor analysis: Preference for Collaborative Learning, Preference for Independent Learning, and Interest and Motivation to Pursue Math. With these statistical controls in place, the ANOVA results showed significant main effects for IBL status and gender, with IBL students and men scoring higher. We also found a gender by IBL group interaction with no difference between women and men in the IBL group, but men scoring higher in the non-IBL group [16].

Finally, among IBL math-track students, first-year students reported higher gains than did later-stage students in mathematical thinking, persistence, and collaboration. Their gains also surpassed their non-IBL peers in these areas as well as confidence and positive attitude about mathematics. A similar pattern was found when data were differentiated by the number of prior college mathematics courses taken, with less experienced IBL students reporting higher gains than more experienced students—a pattern not seen for non-IBL peers. Finally, this pattern was echoed in interviews, where first-year students reported more gains overall than advanced students, especially cognitive gains, changes in their understanding of the nature of mathematics, and affective gains including confidence and enjoyment.

4.2 Attitudes and Beliefs

On the attitudinal survey, among IBL students, most interest and confidence measures remained flat or increased modestly from pre- to post-course, while the confidence of non-IBL students declined (Figure 3). In general, IBL students reported attitudes and beliefs that were more supportive of learning mathematics.

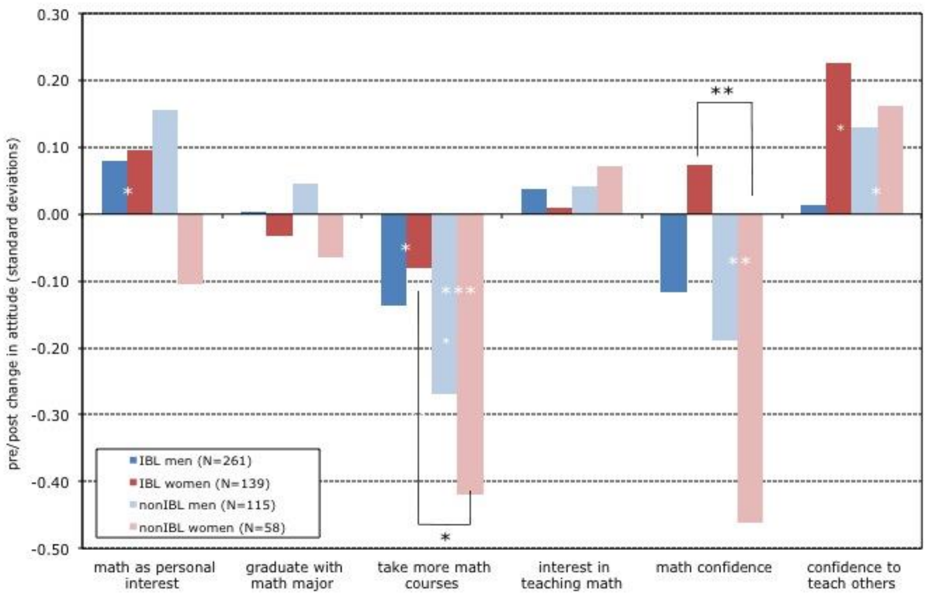

Figure 3: Mean pre/post changes in attitude, for men and women in IBL and non-IBL courses. Items used a 7-point Likert-style scale; the change in standard deviations of pre-survey response is graphed. Asterisks denote statistical significance: *** $p<0.001$; ** $p<0.01$; * $p<0.05$. The pre/post comparison (for any student sub-group) is marked on the bars; between-group comparisons are bracketed and marked outside the bars.

Like the learning gains, pre/post-course changes in students' math confidence and interest differed by gender (Figure 3). Both non-IBL men and women reported losses in their confidence and intent to pursue more math, with women's more pronounced. But after IBL courses, women reported increased confidence to do and to teach math, and greater personal interest in math.

4.3 Grades and Course-taking Patterns

The academic records study examined students' grades and course-taking patterns [17]. The study did not compare grades in IBL and non-IBL courses directly, where assessment tasks and standards may differ, but in later courses, where students who did and did not experience the intervention all enrolled together. These data show that IBL students earned as good or better math grades than their non-IBL peers, and took as many or more later math courses. This was equally true for men and women.

Some learning outcomes also differed among students by prior achievement level [17]. In particular, IBL students with low prior mathematics grades (GPA<2.5) earned the same or better grades in subsequent courses of several types, typically increasing by 0.3-0.4 grade points (on a 4-point scale) (Figure 4). Such improvement was not seen for their low-performing peers in non-IBL courses, nor for their higher-achieving peers, whose grades generally dropped as courses became more difficult.

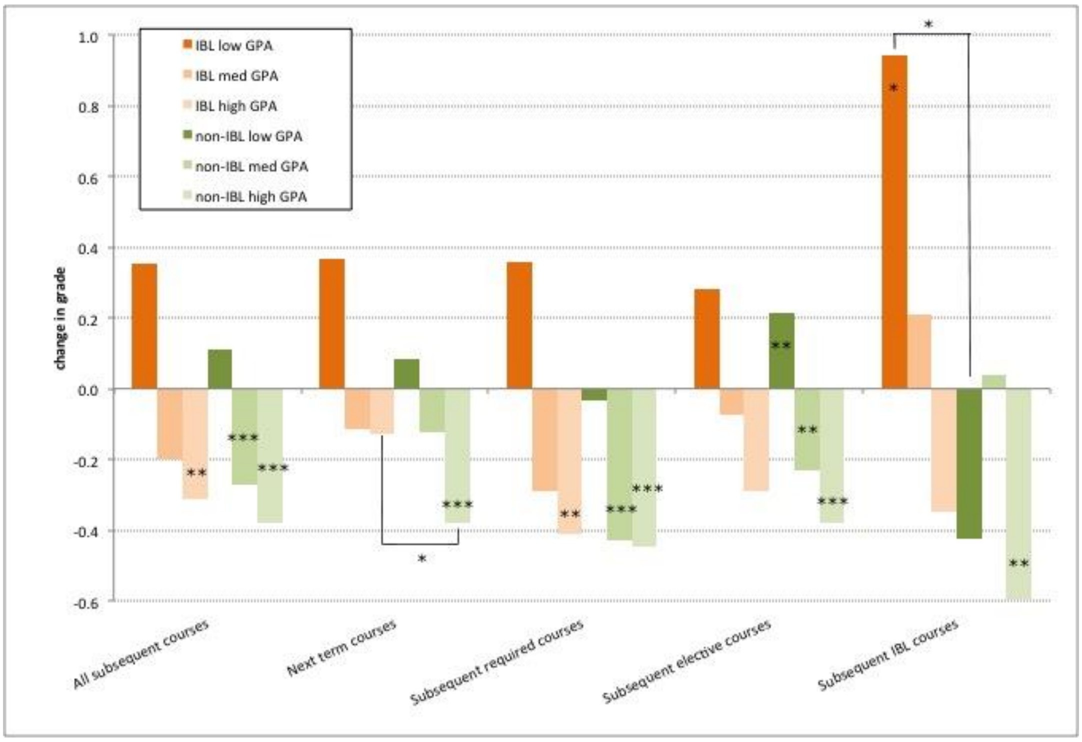

Figure 4: Marginal mean change in average mathematics grade for students who took IBL and non-IBL sections of course L1 in 2001-2008, by prior achievement level. Grades subsequent to course L1, for courses of particular types, are compared with prior math grades using a General Linear Model controlling for the number of prior math courses; see [17] for details. Students are divided by prior grade point average (GPA) on a 4-point scale: Low = GPA ≤ 2.5; Medium = 2.5 < GPA ≤ 3.4; High = GPA > 3.4. Asterisks denote statistical significance: * p<0.05; ** p <0.01; *** p<0.001. Significance of pre-to-post change within a student group is computed from paired

samples test on mean without GLM controls and marked on the bar. Between-group comparisons are based on the GLM analysis and marked by brackets.

4.4 Tests

In two cases, we are able to administer an appropriate content assessment to subsets of students. For pre-service teachers, we used a well-validated external test of learning mathematics for teaching (LMT) [18]. Improved scores on this test have been connected to positive effects on teachers' instruction [19]. Changes in students' pre-to-post LMT scores reflected real gains in understanding, suggesting that IBL courses benefited students preparing to teach in ways that will benefit their future work as teachers [14]. Moreover, test score gains were anti-correlated with initial score: that is, students who had the lowest scores on the pre-test improved most on the post-test. This finding mirrors students' self-report of learning gains, where IBL pre-service teachers with lower overall GPAs reported higher gains than their classmates with higher grades.

A voluntary sample of math-track students took a test of their ability to evaluate mathematical arguments and their reasons for judging arguments to be proofs or not [20]. Both groups performed well; there were no general differences in their scores in total or on specific problems. However, there was some evidence that IBL students were more skilled in recognizing valid and invalid arguments and that they applied more expert-like reasoning in making such evaluations. The small sample of academically strong students limits interpretation of these findings. The test was not appropriate for many courses, thus limiting the utility of content-specific assessments in this setting.

5 Discussion

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 did not. The nature and types of the observed cognitive, affective and social gains were very consistent across multiple data sources, and there is evidence that these gains carry over to benefit students' work in later courses. Thus, despite high variability in the student samples and in the nature of the intervention they experienced, the study detected real differences in student outcomes that align with the literature on the benefits of student-centered approaches, and showed particular benefit for several sub-groups of students.

5.1 Levelling the Playing Field for Women

In the U.S., women's share of undergraduate degrees in mathematics has declined in the past two decades, unlike most other STEM fields [21]. Thus we examined differences in student outcomes by gender. On the SALG-M, women in IBL classes reported learning gains as high or higher than their male classmates across all cognitive, affective and social gains areas, but women in non-IBL classes reported statistically much *lower* gains than male classmates. Thus, while both men and women benefited from IBL courses, traditional teaching approaches did specific disservice to women, inhibiting their learning and reducing their confidence.

The grades data show that, while women reported less mastery and lower confidence at the end of a non-IBL course, their later success was in fact equal to their male peers. This coupled with the sharp attitudinal declines of non-IBL women suggests that the problem with non-IBL courses is primarily affective, not an objective gap in performance. In contrast, IBL courses provided a learning environment that was effective for both men and women in both the short and longer term. Moreover, both men and women persist slightly better after an IBL course [17]—suggesting that better affective experiences may also enhance students' further pursuit of mathematical study.

Student interviews revealed no gender differences in reported gains, and few gender-based differences in their experiences. That gender issues are largely absent in the interview data in fact explains the survey data well: from students' perspective, IBL classrooms offered equitable environments where all could succeed. Indeed, several elements common in IBL courses have been noted as effective for women, including collaborative work [1], problem-solving and communication [22]. Public sharing and critique of mathematical work may foster self-efficacy and link effort, rather than innate talent, to success—enhancing women's sense of belonging to the discipline [23].

5.2 Narrowing the Gap for Lower-Achieving Students

Instructors commonly hypothesized that IBL experiences were most beneficial for students below the top of the class. Thus we examined the data for differences in student outcomes by prior achievement. Two lines of evidence suggest particular benefit for low achievers. First, the LMT test scores for pre-service teachers showed a clear trend that initially low-scoring students improved most on the post-test. Second, analysis of students' grades subsequent to an IBL course showed that, among math-track students, taking an IBL course flattened achievement differences among students. In non-IBL courses, previously low-achieving students gained no ground, but following an IBL course, low achievers' grades were boosted relative both to their own previous performance and to non-IBL peers. No other student groups made such gains.

We surmise that the benefits of IBL to low achievers arise from transferable changes to their problem-solving strategies and study habits. Particularly among students who had not previously developed these skills, this is a powerful effect [24]—yet one without harm to others. Indeed, high achievers were encouraged by an IBL experience to take more mathematics, especially more IBL courses [17]—consistent with instructor observations that strong students found the IBL approach stimulating.

5.3 Jump-Starting First-Year Students

In focus groups, first-year students were particularly enthusiastic about how IBL courses had enhanced their learning in other courses. Thus we examined differences in student outcomes for students taking IBL courses earlier or later in their college careers. Data from surveys and interviews were consistent in suggesting that IBL experiences were more powerful for students earlier in their college career. Positive early experiences may be tied to persistence: after a first-year IBL honors course, students took more mathematics courses than did a matched sample of non-IBL peers [17]. An early experience of IBL may contrast strikingly with students' high school work, and changes in students' approaches to learning or studying mathematics [13] may carry forward to influence their choices and success in later courses.

5.4 Lessons for Studying Change at Scale

So far we have focused on the study results—but there are larger lessons to be learned. One such lesson is that meaningful outcomes of student-centered learning can be measured on a multi-course, multi-institution scale—but evaluating these outcomes is not simple. The setting is inherently messy and the researcher has little control over implementation decisions. Our quasi-experimental research design emphasized general measures, such as self-report and grades, that apply across multiple courses, institutions, and student audiences. Such data provided sufficient statistical power to disaggregate results by student sub-group and to relate student outcomes to demographic and attitudinal covariates; the use of multiple research methods further strengthened interpretation of the results. But such studies are needed to scale up from "proof-of-concept" studies that examine the effectiveness of these instructional approaches in highly controlled but educationally unrealistic conditions.

Second, the study raises questions about whether initiatives like this make any difference, and what they tell us about how to carry out such change. This initiative approaches a scale that can make a substantive difference in student mathematical outcomes and, potentially, their retention in the discipline and workforce. At the four Centers, an estimated 40% of math majors experienced an IBL course during the study period, as did nearly all pre-service K-12 teachers at two sites. Moreover, the study results suggest that, while teachers' skill in implementing IBL does matter, very high fidelity of implementation is not essential to achieving improved student outcomes.

Yet challenges remain for sustaining these efforts. The Centers have made less progress in institutionalizing IBL courses than in changing instructors. Strategies that have worked in some contexts include informing and engaging departmental colleagues and leaders, even those who do not participate in IBL teaching; crafting instructor support through mentoring and team teaching; and using IBL to build a larger community of practice around teaching. Most effective have been efforts to draw in early-career instructors—postdoctoral scholars and graduate students—who come to view IBL as a flexible tool kit for teaching that they can carry along when they move on to new institutions [13]. In future work, we will examine the processes of change that have succeed or failed at the Centers and in the broader IBL movement in mathematics.

Acknowledgements

I thank my collaborators Marja-Liisa Hassi, Marina Kogan and Timothy J. Weston; also Sarah Hough and other E&ER team members. We thank the instructors, students, observers, staff and leaders at the IBL Math Centers, and the Educational Advancement Foundation for its support.

References

[1] L. Springer, M. E. Stanne, and S. Donovan, *Measuring the success of small-group learning in college-level SMET teaching: A meta-analysis,* Review of educational research 69(1) (1999), pp. 21-51.

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

[3] S. R. Singer, N. R. Nielsen, and H. A. Schweingruber, H. A. (Eds.), *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. Washington, DC: National Academies Press, 2012

[4] O. N. Kwon, C. Rasmussen, and K. Allen, *Students' retention of mathematical knowledge and skills in differential equations*. School science and mathematics 105(5) (2005), pp. 1-13.

[5] G. Kuh, *High-impact educational practices: What they are, who has access to them, and why they matter*. Washington, DC: American Association of Colleges and Universities, 2008.

[6] J. Fairweather, *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*, 2008 www7.nationalacademies.org/bose/Fairweather_CommissionedPaper.pdf

[7] C. Henderson and M. H. Dancy, *Increasing the impact and diffusion of STEM education innovations.* A White Paper commissioned for the Characterizing the Impact and Diffusion of Engineering Education Innovations Forum, New Orleans, LA, Feb. 7-8, 2011. http://www.nae.edu/Projects/CASEE/CASEEProjects/26183/36304.aspx

[8] R. L. DeHaan, *The impending revolution in undergraduate science education*. Journal of science education and technology 14(2) (2005), pp. 253-269.

[9] M. T. Hora, *Organizational factors and instructional decision-making: A cognitive perspective*. The review of higher education 35(2) (2012), pp. 207-235.

[10] M. Prince, *Does active learning work? A review of the research*. Journal of engineering education 93(3) (2004), pp. 223-231.

[11] S. Yoshinobu and M. Jones, *An overview of inquiry−based learning in mathematics*. In Wiley encyclopedia of operations research and management science, ed. J. J. Cochran, 2013. DOI: 10.1002/9780470400531.eorms1065

[12] E. Seymour, D. Wiese, A.-B. Hunter, and S. M. Daffinrud, *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, March 2000. http://salgsite.org/docs/SALGPaperPresentationAtACS.pdf

[13] S. Laursen, M.-L. Hassi, M. Kogan, A.-B. Hunter, and T. Weston, *Evaluation of the IBL Mathematics Project: Student and Instructor Outcomes of Inquiry-Based Learning in College Mathematics*. (Report to the Educational Advancement Foundation and the IBL Mathematics Centers) Boulder, CO: University of Colorado, Ethnography & Evaluation Research, 2011. http://www.colorado.edu/eer/research/steminquiry.html

[14] S. L. Laursen, M.-L. Hassi, and S. Hough, *Inquiry-based learning in mathematics content courses for pre-service teachers*. Manuscript in review, 2013.

[15] J. Pearl, *Understanding propensity scores.* In *Causality: Models, Reasoning, and Inference*, 2nd ed. New York: Cambridge University Press, 2009.

[16] S. L. Laursen, M.-L. Hassi, M. Kogan, and T. J. Weston, *Benefits for women and men of inquiry-based learning in college mathematics: A multi-institution study*. Journal of Research in Mathematics Education, accepted, 2013.

[17] M. Kogan and S. L. Laursen, *Assessing long-term effects of inquiry-based learning: A case study from college mathematics*. Innovative higher education, 2013. http://link.springer.com/article/10.1007/s10755-013-9269-9

[18] H. C. Hill, S. G. Schilling, and D. L. Ball, *Developing measures of teachers' mathematics knowledge for teaching*. The elementary school journal 105(1) (2004), pp. 11-30.

[19] H. C. Hill, B. Rowan, and D. L. Ball, *Effects of teachers' mathematical knowledge for teaching on student achievement*. American educational research journal 42(2) (2005), pp. 371-406.

[20] K. Weber, *Mathematics majors' perceptions of conviction, validity, and proof*. Mathematical thinking and learning 12(4) (2010), pp. 306-336.

[21] National Center for Education Statistics (NCES), *Digest of Education Statistics.* US Department of Education, 2009. nces.ed.gov/programs/digest/

[22] X. Du and A. Kolmos, *Increasing the diversity of engineering education—a gender analysis in a PBL context.* European journal of engineering education 34(5) (2009), 425-437.

[23] C. Good, A. Rattan, and C. S. Dweck, *Why do women opt out? Sense of belonging and women's representation in mathematics.* Journal of personality and social psychology 102(4) (2012), 700-717.

[24] M.-L. Hassi and S. L. Laursen, *Transformative learning: Personal empowerment in learning mathematic*s. Manuscript in review, 2013.