Experiencing Calculus Through Computational Labs: Our Department's Cultural Drift Toward Modernizing Mathematics Instruction

December 6, 2019

Abstract: Our traditional model of calculus instruction at a large public research university emphasized factual knowledge and procedural fluency with few realistic applications, leaving a chasm between classroom mathematics and disciplinary practice. Amid an ongoing effort by our department to improve undergraduate learning outcomes, we replaced recitations in Calculus II with computational team labs to bridge this divide. Interview, classroom observation, survey, and gradebook data suggest that the labs facilitated a rich learning experience while student grade outcomes were unchanged. Student engagement, however, were inhibited by institutional structural factors, whose remedy requires a shift in departmental culture concerning the purpose of mathematics instruction.

Keywords: calculus, computational labs, problem of scale, rich learning experience, modeling, qualitative assessment, active learning, group work.

INTRODUCTION

Changing department culture and teaching practices in undergraduate mathematics courses is a complex process, resisted by the inertia of the status quo. This paper describes one step of this process at a large public research university: incorporating computational labs into mainstream Calculus II instruction. These labs served to connect calculus concepts to disciplinary STEM practice by asking students to analyze and interpret the results of numerical simulations, expose students to computational modeling, and provide students with a team-based learning experience. Through prepared MATLAB programming notebooks, students investigate contexts such as rocket science, disease modelling, and market economics forecasting by modifying and executing provided code, guided by both factual and interpretive questions concerning the results of their simulations and the underlying mathematical models. We also report on our qualitative educational research designed to evaluate the effectiveness of the labs and inform ongoing development. Student and instructor interviews, classroom observations, surveys, and grade data were collected to understand the impact on student outcomes and the factors that shape student engagement. This project reflects an ongoing effort to improve undergraduate teaching in our department: shifting curriculum towards rich, collaborative experiences with interdisciplinary contexts and away from transferring and assessing individual mathematical domain knowledge. Notably, improving grades and DFW rates was *not* one of the design goals because the historical distribution was acceptable to the department prior to the reform. Instead, this reform focused on reimagining students' learning experience and promoting higher-level, interdisciplinary learning goals.

In this paper, we first describe the factors that provided the impetus for the reform and explain the context. Second, we explain the details of the implementation and the steps to scale-up the reform. Third, we present our educational research, including the methods for early feedback and formal data-collection that took place later. We conclude with a summary of the lessons learned and some advice for fellow reformers.

1 IMPETUS FOR REFORM

A confluence of several factors provided the impetus for our department to commit significant resources to begin a reform of the calculus sequence. First, there was a growing recognition among many of our faculty that a *traditional model* of instruction—which we take to mean a lecture-style delivery, an emphasis on knowledge transfer and rote calculations, and a reliance on high-stakes, individual, summative assessment—failed to meet the learning needs of many students. In this model, students are expected to acquire, either heuristically through practice or by explicit instruction, a set of procedures for solving problems in a limited context; the procedures ranging from technical ("how to perform partial fraction decomposition") to decision-making ("how to choose which series convergence test to apply"). Success is demonstrated through correct application of the procedures to the same limited context. A traditional instructional sequence might include: (a) a classroom demonstration of integration by parts, (b) an assignment of 10 exercises that require integration by parts, and (c) a quiz with a similar problem. The content most compatible with this instructional model is precisely the computational fluency that can be made obsolete by technology and is de-emphasized in recent policy recommendations such as the Curriculum Foundations Project [6, 7]. The learning goals in such a course are typically limited to lower-level cognition: *remembering* and *applying* in the revised Bloom's taxonomy [1] or *memorization* and *procedures without connections* in the Mathematics Tasks Framework [11]. Furthermore, active learning has been shown to improve outcomes, especially for women and traditionally underrepresented minority students [5], although recent studies call into question the generalizability of earlier findings [8].

Second, our department received specific feedback from the Engineering College about failures of the traditional calculus sequence in preparing students for upper-division Engineering courses. A key observation was that students struggled with calculus applications in their Engineering classes, thought to be caused by a perceived disconnect between calculus content and the core engineering curriculum that led to low student engagement, as reported retrospectively by upper-division Engineering students. This can be attributed both to the typical temporal gap between students' enrollment in calculus classes and the recall of calculus concepts in their major courses, and to the limited scope of applications in the traditional model. An emblematic example is that of the "sequences and series" concepts of Calculus II: Engineering students typically enroll in Calculus II during their first three semesters, but these concepts do not resurface in their major courses until junior- or senior-level courses when Fourier series or numerical methods are discussed. The emphasis on the procedural application of series convergence tests, rather than the applications of the concepts as powerful tools in approximating solutions in the applied sciences, compounds the engagement issue. The feedback from the Engineering College provided convincing evidence to the rest of the faculty of the need for action.

Third, the university's support of interdisciplinary research programs around the applications of computational mathematics in natural sciences and engineering is driving the inclusion of computational mathematics into undergraduate curricula. The labs were designed to promote and demystify the integral role of coding in modern applications of mathematics. The computational environment allowed us to focus the instruction on higher-level tasks by offloading the procedural aspects of calculus to the computer. It also allowed us to explore more interesting problems and showcase the applications of calculus in more realistic contexts.

2 LOCAL CONTEXT AND DEPARTMENT CULTURE

Our institution is a large public research university serving over 40,000 undergraduate students. Calculus II serves approximately 2000 students annually. Because of the scale, Calculus II is administered as a uniform course to facilitate comparable student experiences across the approximately 40 sections per semester; a common syllabus enforces uniformity of the schedule, assessments (common midterm and final examinations accounting for 70% of student grade), and online homework (10% of student grade). Weekly instruction consists of three 50-minute lectures (150-250 students; faculty lecturer) and one 50-minute recitation (30 students; graduate teaching assistant). A traditional recitation typically included 10-15 minutes of student questions, 15-25 minutes of demonstrated examples (selected by the instructor), and a proctored individual quiz. Some recitation leaders incorporate group activities, but the more common approach is teacher-led demonstration with some whole-class discussion.

Teaching faculty, research faculty, and post-docs teach large-lectures that serve most of the students during fall and spring semesters; each large lecture is divided into 7-10 recitations. The standard load for teaching faculty is 6 courses per year. Research faculty teach 3 course per year, with equal division of time between teaching and research and a smaller portion devoted to departmental service. Support for instruction and organization is provided by the Calculus Coordinator, who helps maintain continuity and uniformity between semesters, handles the administrative aspects (e.g. student complaints and academic integrity issues), and works with the lecturers on creating the uniform exams. The department also promotes teaching excellence through the Center for Instructional Mentoring, which coordinates mandatory and elective professional development for graduate students, postdocs, and new faculty via classroom observations, teaching orientation, and peer mentoring. Anecdotally, these efforts are well-received and the departmental culture about teaching has improved since its inception in 2016.

There have been several similar reform efforts within our department, which is an indicator of a growing departmental commitment to improving students' learning experiences across the board. One set of reforms involved separating the traditional college algebra class into two pathways that served different populations—students seeking to satisfy graduation requirements and students pursuing STEM degrees—more effectively. A quantitative literacy sequence was developed for students who need to satisfy general education requirements, in lieu of college algebra. A two-semester college algebra sequence was developed to replace the traditional sequence of Intermediate Algebra (remedial-non-credit-bearing) and College Algebra, for students who need algebra as a prerequisite for statistics or Survey of Calculus (a terminal course for Life Science and Business students). Perhaps a precursor to the Calculus II reform, our Survey of Calculus course was reformed to include group-based application labs that explored calculus applications in biology and business contexts. Following the start of the Calculus II reform, similar computation lab-type activities were developed for Differential Equations.

3 INITIAL DESIGN AND IMPLEMENTATION

Our initial implementation plan was designed as a three-year project. The focus of the first year was creating and piloting the lab activities in a small class (30 students/semester) co-taught by the two tenure-stream faculty leading the project. The department took the unusual step of double-staffing a small class in lieu of the usual one-faculty-per-large-lecture structure to enable more direct in person feedback on the progress of the development effort. Additional teaching releases were also awarded to the two early-career research faculty who committed to developing the labs from scratch. The lab development was based both on past teaching experience of the faculty leads, as well as on topical suggestions resulting from discussions held by the two faculty leads with faculty representatives from each Engineering department of the university. In the second year, the faculty leads co-taught medium-sized lectures (80 students/semester) divided into two recitation sections in which the labs were led by a pair of teaching assistants; development focused on revising the lab materials and addressing logistical issues with scaling up the labs. In the third year, a large-scale pilot was implemented in which the faculty leads co-taught one full-sized lecture (160 students/semester) and the labs were used in all 7 recitation sections, each taught by a single teaching assistant (the regular structure); development focused on revising the labs for at-scale deployment and developing teaching supports for lab instructors.

In the fourth year of the reform (the 2019-2020 academic year), the labs were being run at-scale. The labs were being used in all 24 sections in the fall semester (across 4 large-lectures), during which the faculty project leaders were both on leave. The labs were implemented in all 38 sections in the spring semester (across 5 large-lectures), with one of the faculty leads stepping back in as a lecturer and the course supervisor.

3.1 Rationale for Labs Contained in Recitations

The labs were designed as replacements of traditional recitations to maintain coherency in the curriculum throughout a gradual change, make the reform sustainable, and address logistical challenges of large-scale teaching. This section explains four factors that led to this design approach.

First and foremost, our Engineering colleagues expressed strong desires for the mathematical content of the courses to remain unchanged, which aligns with mathematics faculty disposition. Additionally, core content changes would necessitate a re-examination of the entire calculus sequence to ensure that content arcs remained intact. Therefore, the lecture content was not modified.

Second, compartmentalizing the reform within recitations provides some independence from the instructional decisions of individual lecturers. Logistically, any instructor in the department can teach the large-lectures without involvement in the development or implementation of the labs because teaching support for the lab instructors is provided departmentally through the Center for Instructional Mentoring. Locating the labs within recitation also allows faculty instructional autonomy in the lectures, which helps soften potential pushback against mandated instructional reform.

Third, the structural independence of recitations facilitated the gradual roll-out and assessment of the labs by allowing the parallel running of pilot and traditional sections. In the pilot sections, we were able to run extended-time recitations in the first two years (120 minutes and 80 minutes respectively) to hone the lab material and discover the challenges faced by recitation instructors with diverse teaching experience and skills. These findings informed the revision of lab material and the development of the teaching support required to deploy the labs at scale. For the third-year pilot, we returned to 50-minute recitations to stay aligned with university credit-hour guidelines.

Fourth, the labs were designed to fit within the traditional grading framework. The portion of students' grades allotted for quizzes in the traditional model (20%) was split between revised quiz-type assignments aimed at self-assessment (14%) and the labs (6%); the emphasis on uniform exams was maintained (70%). We designed the labs to be contained within the recitation without extra out-of-class work, in view of our gradual roll-out. Previous studies reported reform efforts derailed by comparisons between traditional and reformed sections in which students determined that the reformed sections required more work for the same grade [4].

3.2 Engineering Focus

The enrollment in these pilot courses was initially planned to be restricted to Engineering students, with the long-term target of deploying computational labs only in special sections running in parallel to traditional sections. The restriction of enrollment to Engineering students came from two practical considerations. First, the Engineering College requested that labs be designed using the MATLAB language and provided access to the JupyterHub service that was initially used to provide cloud access. Cloud computing access was deemed essential to minimize the technological barrier: students could run the labs on personal computing devices without installing any software. Second, we anticipated collaborating with Engineering colleagues to build labs around disciplinary topics and wished to avoid marginalizing non-Engineering students with unfamiliar contexts.

During the third year of the reform, the Engineering enrollment restriction was lifted both because the labs were transitioned to MATLAB LiveScripts (available to all students through our university site license, including cloud access) and because we were able to reformulate contextual applications to be more broadly accessible to all students. For example, Lab 2 is an exploration of rocketry, which is plausibly interesting to a wide range of students outside of Engineering (see Appendix D for descriptions of each lab). As the labs developed, it became clear that the activities had the opportunity to be beneficial for all Calculus II students, and the labs represented a shift towards adopting higher-level and interdisciplinary learning goals in the calculus curriculum in general.

4 REFORM ASSESSMENT PLAN

This reform is the first in our department to target the mainstream calculus sequence, which is the most resource-intensive and most visible sequence (compared to the quantitative literacy, college algebra and statistics, or applied calculus for business and the life sciences). As such, systematically evaluating its effectiveness is crucial to inform revisions and facilitate adoption by the wider teaching community.

4.1 Years 1 and 2 – Early Feedback

Feedback was collected during the first two years of the pilot to inform ongoing revisions. During the first year, students were surveyed after each lab to provide feedback on logistical elements (i.e. time required), the difficulty of the tasks, and group dynamics (see Appendix A). During the second year, feedback was solicited from lab instructors both before and after the labs to identify aspects of the labs that were challenging for students. At this phase, feedback and revisions focused on adjusting the content and nature of the tasks in the labs to improve students' experiences. To enable the transition to shorter classes, the labs were shortened by focusing the investigation on one aspect of the application contexts in detail, rather than exploring multiple facets. The computer code was also reorganized to be more intuitive (less computationally optimized, but easier to interpret) and thus more suitable for a broader student population.

4.2 Year 3 – Qualitative Educational Research

During the third year of the project, we conducted qualitative research to evaluate the effectiveness of the labs by examining student engagement along behavioral (the observable actions that students do), emotional (affective components of engagement: involves students' emotional reactions to a task such as interest or curiosity, the importance of doing well, the relationship between the task and students' learning and personal goals, and the cost of completing the task), and cognitive (the psychological investment that students make to complete a task in a certain way, including thoughtfulness and willingness to exert effort to comprehend the learning goals of the task) dimensions [10]. Learning occurs through engagement with academic tasks, thus understanding the nature of student engagement is a lens to understand students' learning experiences [2]. Framing the study this way is important because gains on traditional achievement measures (e.g. exams, course grades) were neither expected nor targeted by design. The uniform course exams focus on individual demonstration factual knowledge and procedural fluency, whereas the labs focus on collaborative problem solving and modeling of realistic problems. By examining student engagement, we uncovered students' learning experiences with the labs independent of the uniform course exams (although their experience was impacted by the uniform structure).

Data collected included interviews, classroom observations, and a post-course survey, and grade distribution data. We interviewed 30 students using a semi-structured protocol (see Appendix B) designed to collect background information (dispositions towards mathematics, history with mathematics courses, goals for Calculus II) and engagement data (activity during the lab, group dynamics, experiences with MATLAB, perceptions about the extent to which the labs facilitated learning, and investment in completing the lab). We also conducted 20 video-recorded classroom observations each capturing one team's experience. All calculus students are asked to respond to pre/post-course surveys, and additional survey items were added for students in lab sections (see Appendix C). Grade data is routinely collected by the department, along with detailed enrollment data.

5 RESEARCH FINDINGS

The labs present an *opportunity* for students to engage with realistic mathematical modeling in contexts that many students find interesting, but this opportunity was unrealized for many students, with the survey data painting a more negative picture than the interview data. A majority of the interview participants reported that the labs exposed them to interesting applications of calculus, that they enjoyed the variety that the labs provided, that they appreciated working on challenging problems, and that they found the experience using MATLAB to be useful. Illustrating how the labs can help clarify concepts, one student exclaimed during a lab, "Oh, that's what convergence means!" In this sense, the labs are a useful teaching tool for inspiring curiosity, and several interview participants expressed that they were frustrated when they could not finish or fully grasp the labs because they wanted to "figure it out". Furthermore, one interview participant reported that he enrolled in a computational modeling course after his experience with the labs, in spite of his overall struggle with Calculus II; this is an inspiring finding indicating that the labs might help to counteract the switcher-persister imbalance that leads to many students leaving STEM fields after negative experiences in calculus courses [3].

Students continued to put forward effort to complete the labs throughout the semester, but had negative feelings at the end of term, which was especially apparent in the post-course survey data. One participant explained concisely, "The idea of the labs is great, but execution is failing." Students interviewed before the mid-term generally had a positive disposition towards the labs; they viewed the labs as interesting, enjoyable, and appreciated the instructional variety. The post-course survey and follow-up interviews, however, painted a very different picture. Participants interviewed near the end of the term or after the final exam commented that, while the labs were interesting, they ultimately did not impact their success in the course (referring to grades), as could be expected with the final exam not focusing on the lab content. The survey results were more bleak, but they mirror overall negative shift in student dispositions near the end of the semester:

[Insert Table 1]

[Caption] Table 1: Responses from Year 3 post-course survey. (n=92)

Interview participants were evenly split about their motivation for completing the labs as the semester progressed: some found them interesting and appreciated the challenge, while others completed them simply to earn course credit. We interpret this data to mean that behavior engagement was maintained, but emotional and cognitive engagement declined throughout the semester. While these engagement issues are significant, we believe they will be addressed by further curricular commitment to the labs; in other words, rebalancing grade distribution to make the lab grades more impactful and including a final assessment of the lab content.

5.1 Labs Within a Traditional Curriculum

As previously described, the labs were designed to replace traditional recitations without major reforms to the rest of the course. This structure contributed to several issues. First, most interview participants reported that the labs did not feel integral to the course and their role was unclear: the labs were in a different classroom, with a different instructor, were not discussed during lecture, and the content was not included on exams. The most effective lab instructors reiterated that the purpose of the labs was to explore realistic applications of calculus and explained the specific application of each lab at the beginning of class. This effort helped their students understand the purpose of each lab and improved engagement. Future professional development will focus on preparing lab instructors for short whole-class discussions connecting lab contexts to calculus content.

Second, nearly all interview participants reported that they felt rushed to complete the labs, which interfered with their cognitive engagement. Classroom observations and interviews revealed that the major contributor to the length issue is students needing 10-20 minutes to read the lab introduction and understand the contextual problem at the start of recitation. Several interview participants expressed their frustration that they were not able to complete the lab in the allotted time and requested an opportunity to think about the lab ahead of recitation. To address this issue, we split the labs into a *prelab* portion—a short video and a short set of questions to be completed individually before class—and a *lab* portion that contains the activities that benefit most from group discussions with multiple points of view and diverse expertise; continuing inquiry will investigate the effectiveness of this adjustment.

Third, engagement was hindered by students' perceptions that the labs had little impact on their grades. Calculus II students are generally dedicated to their academic success, toward which they accurately determined that uniform exams were the most important component. During an interview in the middle of the semester, one participant reported that the labs were enjoyable and interesting, but during a post-course interview opined that the labs were extra work that did not help her earn a better grade, a sharp shift in perception. This is exacerbated by the uniform exams (by design) not including lab-specific material. Future iteration of the labs will include an adjusted grading scheme with an increased emphasis on the labs, including a group-assessment element at the end of the class (a lab practicum).

5.2 Groupwork

Overall, the group-based lab experience was positive for most students. One concern among mathematics faculty was that group-based assignments will burden students who may perceive that they are disadvantaged by weak group members, especially when a group grade is assigned, which we found to be an unfounded concern. Students overwhelmingly reviewed the group setting as a positive learning experience, especially when the diverse expertise within the group structure provided support when they struggled to grasp the conceptual mathematical ideas or lacked the programming experience to interpret and edit the MATLAB code. On the other hand, structural adjustments and more professional development for lab section leaders is required to foster the development of cohesive groups. One specific issue arose from the random assignment of students to groups at the beginning of lab. This practice was chosen to alleviate logistical issues, namely that of absent or tardy students leaving their group depleted. Survey responses indicated that students preferred being assigned to groups rather than choosing their own group, supporting elements of this practice. However, interview participants indicated a lack of commitment to building relationships with their group members because of the weekly shuffle. Several students suggested that the practice be restructured to facilitate better group development because they wanted to foster a working relationship with their group members over several weeks. Students suggested this would have both logistical advantages and would help them learn more effectively because less time would be spent meeting-and-greeting and there would be more peer accountability knowing that you will be working with the same people next week. Going forward, we will provide lab instructors with support to assign effective groups based on classroom interactions to promote cohesive groups that will work on several labs together before shuffling., a new practice for essentially every lab instructor.

5.3 MATLAB

Both interview participants and the lab section leaders reported that the computational components became more familiar throughout the semester and that most students could operate within the MATLAB environment. Several students explained that they enjoyed the labs because it helped them see how mathematics might "actually be used by engineers". Many students expressed concerns with the coding component at the beginning of the course (they were unfamiliar with MATLAB, felt lost during the first lab, and wished for a supplementary MATLAB "crash-course" to help them with the coding), but most of these concerns are alleviated by the end of the course (participants explained that they were able to learn the coding skills required to complete the labs, were usually paired with at least one group member who was familiar with MATLAB, and realized that minimal original coding was required to complete most labs). Because the negative initial reactions to MATLAB subside as students become more comfortable with the coding throughout the semester, plans to offer coding support at the beginning of the semester are on hold to allocate development resources to other areas of need.

6 CONCLUSION

Our findings have provided important guidance to further our ongoing reform and are likely to inform similar efforts. The research has also generated some lessons that can be generally applied to reform efforts in other contexts.

6.1 Pedagogical Inertia

We should expect that the initial implementation of reform efforts will be problematic and have more logistical issues than the traditional model, due to pedagogical inertia: traditional curriculum elements have been "polished" over many years and practice has evolved slowly so that large courses run smoothly. Extensive structures have spawned to support traditional teaching practices, and their analogues have not yet developed for reformed teaching. Some examples include: institutional reliance on heavily-weighted, individual, timed exams as cheap (they can be proctored and graded by undergraduate hourly hires) and clean (they produce precise one-dimensional numerical scores) assessment metrics; existence of a well-trained labor force (current instructors are often themselves trained on the traditional model); and student expectations of mathematics being a subject in which "there is one right answer" and mastery of rote procedures leads to success.

Long-term reform requires replicating different versions of these structures for reformed teaching methods. Interview participants reported the connection between their efforts on the labs and achieving a high course grade was unclear; the course grading scheme needs to be rebalanced to align with the instructional emphasis (less weight on exams, more weight on labs). Additionally, professional development programs must be developed to train instructors to facilitate group work. Perhaps the most challenging hurdle is adjusting student expectations: that mathematics courses should involve higher level cognitive skills and perseverance through complex problems, instead of routine calculations solved using rote procedures.

6.2 Paradigm Shift

A paradigm shift is required to propel the adoption of reformed teaching practices, particularly regarding assessment practice, and this can be done without immediately rejecting traditional philosophies. Resistance to the full deployment of the labs emerged from a department culture that equates evidence of student learning to individual performance on timed, proctored, summative assessment, where a passing grade is assumed to certify the acquisition of some amount of "knowledge". These metrics do not and cannot account for students' ability to effectively employ resources, work effectively with a team, or creatively solve new problems. An alternative assessment philosophy assumes that learning emerges from social interactions that are irreducibly complex, and the resulting knowledge and growth are not well-suited to be measured on individual, timed exams. From this perspective, we expect each student to contribute unique knowledge and skills, leading to social knowledge that is greater than the sum of its parts, and group members draw from the socially constructed knowledge to create their own meaning.

Interview participants reported frustration that it was difficult to implement a strategy to earn high score on the labs, related to the practice of grading the labs based on accurate completion of the entire lab. Students who faced challenges with coding for the first time or lacking contextual knowledge useful for understanding the labs (e.g. physics or engineering) often struggled to complete the labs within the class period. Each student has a unique individual experience, so it is unreasonable to expect uniform outcomes and impossible to certify uniform learning. Instead, we must work to ensure that each student has a rich experience that fosters growth, with diverse modes of engagement and opportunities to contribute. Group-worthy tasks—i.e. those which require complex problem solving, provide students with multiple entry points and opportunities to demonstrate competence, deal with authentic and discipline-based content, and require positive interdependence as well as individual accountability—cannot be reduced to individual tasks because the collaborative element is foundational [9]. In this sense, a passing grade certifies that a student has engaged in a sufficient number of rich experiences that require understanding of the connections between mathematical ideas, collaboration with group members, and creative problem-solving. Grading practices need to be developed (and adopted) that address the differences between a timed, narrow-scope, individual quiz and a lengthy, complex, group task.

The slow march away from high-staked individual exams to multi-modal assessment that includes group-grades for rich group tasks is one aspect of the paradigm shift that is required for lab-type activities to fit sustainably within mainstream calculus curricula. Our finding that exam grade distributions were approximately equivalent in the lab and traditional sections helped open the discussion on this paradigm shift: the "do-no-harm" evidence assures skeptical colleagues that students are still learning the traditional content matter, while also engaging in rich computational modelling. As result, the implementation of the labs has continued to expand, and all Calculus II sections used the labs in the fourth year of the project.

6.3 Challenges of Piloting and Evaluating Reform Simultaneously

It is important to recognize the challenges of piloting a reform while simultaneously evaluating its effectiveness, which we suggest necessitates qualitative inquiry.

The first challenge is that comparisons of student achievement data are weak if the reformed curriculum emphasizes new learning goals. The traditional assessments do not capture the new learning goals and thus comparisons are limited to "do-no-harm" findings, which can still be important (as described previously). We cannot use this traditional achievement data, however, to quantify the learning gains from the labs. Furthermore, new learning goals are likely to evolve throughout a reform, so it is difficult to stabilize quantitative assessments prior to implementation, leaving qualitative analysis as the best option for understanding student engagement with the reformed curriculum.

The second challenge is that a gradual scale-up of the reform necessitates that pilot reformed sections are run in parallel to traditional sections, which introduces implementation problems that can deteriorate student engagement [4]. We found that students compared their experiences across sections and those in the lab sections often determined that they were doing extra work for little reward (other than learning). With the end of the semester in sight, the pressure of high-stakes, uniform examinations overwhelmed students' initial positive perceptions of the labs as they (correctly) identified that doing well on their exams was the most important step towards pursuing their goals (e.g. being admitted to the College of Engineering). While there is no evidence that the labs systematically disadvantaged students on the uniform exams, it is nearly impossible to convince students that any content is important if it is not on the exam. In other words, "what you test is what you get" [12]. As a result, we should expect negative affective responses in post-course survey data, which was exactly what we found, and we must rely on qualitative data to understand the nuance of students' experience.

The third challenge is that small implementation glitches can undermine student engagement and negatively impact students' learning experiences to the point where the reformed element is rejected by many students. Small adjustments to in-class timing, grading schemes, communication of expectations, technology configuration, and instructor preparation can completely change students' experience, to the point that quantitative measures have little meaning until the instructional structure is stabilized. Because of this, we advise caution regarding achievement measures (exam grades, DFW rates) and affective measures (surveys) during the first several iterations of a reform.

In light of these considerations, we advocate for qualitative investigations of student engagement, namely student interviews and classroom observations, throughout the implementation of a reform as a reliable method to understand the extent to which the reform facilitated the intended learning experience. This approach empowered the faculty involved in our reform to draw on their teaching experience and expertise to understand the nuances of students' learning experiences to make appropriate adjustments as we continue the process of reforming mainstream undergraduate mathematics education.

Bibliography

[1] Anderson, L., Krathwohl, D., & Bloom, B. 2001. *A Taxonomy for Learning, Teaching, and Assessing: a Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

[2] Doyle, W., & Carter, K. 1984. Academic tasks in classrooms. *Curriculum Inquiry*. 14(2): 129–149.

[3] Ellis, J., Kelton, M., & Rasmussen, C. 2014. Student perceptions of pedagogy and persistence in calculus. *ZDM - The International Journal on Mathematics Education*. 2(1): 1–20.

[4] Flint, D., Hunter, B., & Kemp, D. (2011). South Dakota State University's participation in the MAA Renewed College Algebra Project. In S. Ganter & W. Haver (Eds.), *Partner Discipline Recommendations for Introductory College Mathematics and the Implications for College Algebra* (pp. 73–76). The Mathematical Association of America.

[5] Freeman, S., Eddy, S., McDonough, M., Smith, M., Okoroafor, N., Jordt, H., & Wenderoth, M. 2014. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23): 8410–8415.

[6] Ganter, S., & Barker, W., (Eds.). 2004. *The Curriculum Foundations Project: Voices of the Partner Disciplines*. Washington, DC: Mathematical Association of America.

[7] Ganter, S., & Haver, W. (Eds.). 2011. *Partner Discipline Recommendations for Introductory College Mathematics and the Implications for College Algebra*. Washington, DC: Mathematical Association of America.

[8] Johnson, E., Keller, R., Andrews-larson, C., Fortune, N., & Keene, K. 2018. Inquiry and inequity in the undergraduate mathematics classroom. *Proceedings of the 40th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Greenville, SC: University of Southern Carolina & Clemson University.

[9] Lotan, R. Group-worthy tasks. *Educational Leadership*, 60(6): 72-75.

[10] Middleton, J. A., Jansen, A., & Goldin, G. A. 2017. The complexities of mathematical engagement: Motivation, affect, and social interactions. In Cai, J. (Eds.), *Compendium for research in mathematics education* (pp. 667–699). Reston, VA: National Council of Teachers of Mathematics.

[11] Stein, M., & Smith, M. 1998. Mathematical tasks as a framework for reflection: Research to practice. *Mathematics Teaching in the Middle School*, 3(4): 268–275.

[12] Wiggins, G. 1988. *Educative assessment: Designing assessments to inform and improve student performance.* San Fransisco, CA: Jossey-Bass Publishers,

APPENDIX A: YEAR 1 LAB SURVEY QUESTIONS

During year one, student teams are asked to answer the following three questions for each lab.

- Which question on this lab did your group spend the most time on?
- Which question on this lab did your group find most difficult? Why?
- Was there a question on which members of your group disagreed with each other on the answer? Briefly describe the different proposed answers, and how your group settled on the answer given in the lab report.

APPENDIX B: INTERVIEW PROTOCOL

This interview protocol is designed to guide a semi-structured interview about the participant's engagement with a group-based lab investigation. Background and demographic data are covered first, followed by questions about specific elements of the labs.

Background

1. How is your calculus class going? What is going well? What things stand out that aren't as helpful or hold you back? How do you feel about your instructor?
2. Please tell me your current academic class and list the college mathematics classes you have taken previously. AP calculus? Repeat courses? Courses at different institutions?
3. What is your major and what is your planned career trajectory?

Disposition toward mathematics

4. What are your feelings about mathematics in general? Skill in mathematics? Confidence with mathematical ability?
5. How does your mathematics class make you feel? Do you feel like you can succeed in the class? Do you feel like you belong in the class?
6. How has MTH 133 informed your view about the role of mathematics in your current career goal?

Goals and Study Strategy

7. How did you decide to enroll in MTH 133? (requirement, you think it will be useful)
8. What are your goals for MTH 133?
9. Can you explain a typical week of your learning associated with MTH 133? I am trying to get a picture of how you learn calculus, so tell me any details that you think are relevant.
    a. What do you do during lecture? During recitation?

b. What do you do outside of class?
c. How do you prepare for quizzes or exams?
d. Do you visit the MLC, office hours, or some other kind of out-of-class instruction?
e. What kind of homework do you do? Webwork? Other problems?

About the Labs

10. Can you explain what typically happens or **what you typically do** during your recitation/lab?

   a. What do you do? Solve a problem using pen and paper? Writing/editing MATLAB code? Analyzed/interpreted results?
   b. What about your group members? What about other classmates/groups?
   c. What does your instructor do?
   d. Does the lecturer mention/launch/debrief the labs during lecture?

11. How does **MATLAB** impact your learning experience?

   a. Did you gain insight about calculus concepts by seeing how they are represented in computer simulations? If so, which concepts and what insights?
   b. Do you feel proficient with MATLAB? Modifying existing code? Writing original code?

12. Do you use any outside **resources** (other than MATLAB) while you are completing your labs? (How do you use them / what do you use them for?)

   a. Calculators? Websites? Classmates in other sections?

13. Do you feel like the labs **help you learn**?

   a. What do the labs help you learn, specifically?
      i. Practice/preparation for exams/quizzes?
      ii. Understand connections between concepts?
      iii. Understand/appreciate applications of calculus ideas?

iv. Practice modelling?

b. Do you feel like class time spend on the labs is worthwhile in the context of succeeding in the course?

i. Is there something else that would be a better use of class time to help you succeed in the course? (Perhaps compare the goals of the labs and quizzes and compare the effectiveness of those goals.)

c. Do you feel like class time spend on the labs is worthwhile in the context of your long-term goals at MSU?

i. Is there something else that would be a better use of class time to help you achieve your long-term goals?

14. How do the labs **make you feel** about your learning?

a. Are you interested in the labs?

b. Do you find the labs enjoyable? What aspects?

i. Working in groups?

ii. Doing something different than solving problems?

iii. Working through problems without the instructor?

iv. Seeing how calculus can be used in practical contexts?

15. Are you **invested** in the labs?

a. Why do you complete the labs? Trying to finish them? Good learning opportunity? Interesting?

b. If the labs were ungraded, would you attend recitation to complete the labs?

16. Are there any changes that you would make to the way that the labs are done?

a. Introduction/follow up?

b. Related homework assignment?

c. Tied to quizzes/exams?

APPENDIX C: SURVEY

The survey is administered as a pre and post-course survey using Qualtrics and a link provided on the course website. Students are informed about the survey by an email from the course supervisor and are also reminded by their instructors during class. Students will receive extra credit for completing the survey (as is typical for the course), and are asked to consent to their survey responses being used as research data.

A. Intro Page

About this Survey

This survey will ask you about your mathematical experiences in calculus, your perceptions of math, and your time commitments this semester. This survey is confidential. Your name and email will be used by a departmental administrator so that we can identify who has taken the
survey. Your instructor may request the results of the survey but all personally identifiable information will be removed prior to sharing the results with them. Accurate data is very important to us. Future departmental decisions and policies may be based on the results of this survey. Therefore we ask that you please electronically sign the statement below:

I will answer the following survey questions truthfully and to the best of my knowledge.

[Electronic Signature box]

B. Consent for Research Participation

Separate from the request to complete this survey for course improvement, you are also being asked to participate in this survey for research. Researchers are required to provide a consent form to inform you about the research study, to convey that participation is voluntary, to explain risks and benefits of participation, and to empower you to make an informed decision. You should feel free to ask the researchers any questions you may have.

Study Title: Student Engagement with Labs in MTH 133
Researcher: Andrew Krause
Email: krausea3@msu.edu
IRB #: STUDY00001288

Check the box below to indicate that you agree to participate in a study about students' experiences in MTH 133. By agreeing to participate in the study, you verify that you are at least 18 years of age. Your participation in the study entails completion of this survey. Participation in this research project is completely voluntary. You have the right to say no. You may change your mind at any time and withdraw. You may choose not to answer specific questions or to stop participating at any time. Whether you choose to participate or not will have no effect on your grade or evaluation. The study will have no cost.

Do you agree to participate in this study? [Yes/No]
If you have questions or concerns about your role and rights as a research participant, would like to obtain information or offer input, or would like to register a complaint about this study, you may contact, anonymously if you wish, the Michigan State University's Human Research Protection Program at 517-355-2180, Fax 517-432-4503, or e-mail irb@msu.edu or regular mail at 4000 Collins Road, Suite 136, Lansing, MI 48910.

Written Responses

1. [pre-survey only] Why did you decide to enroll in MTH 133? Describe the factors that influenced your decision.
2. [post-survey only] How has MTH 133 informed your view about the role of mathematics in your current career goal?

Student Status

3. For which MTH Course are you filling out this survey?
    a. MTH 132
    b. MTH 133
    c. MTH 234
4. What Class level do you consider yourself?
    a. Freshman
    b. Sophomore
    c. Junior
    d. Senior
    e. Other [fill-in-the-blank]

Career Goals

5. Which of the following BEST describes your current career goal?
    a. a STEM career (including health and social sciences)
    b. a career in education
    c. a career in other fields
    d. undecided

[if (a) a STEM career]

6. Which of the following BEST describes your current career goal?
    a. Medical professional (e.g., doctor, dentist, vet.)
    b. Other health professional (e.g., nurse, medical technician)
    c. Life scientist (e.g., biologist, medical researcher)
    d. Earth/Environmental scientist (e.g., geologist, meteorologist)
    e. Physical Scientist (e.g., chemist, physicist, astronomer)
    f. Engineer: Electrical, Computer Science
    g. Engineer: Civil, Environmental, Biosystems/Agriculture
    h. Engineer: Mechanical and Aerospace
    i. Engineer: Chemical and Materials
    j. Mathematician
    k. Social Scientist (e.g., psychologist, sociologist)
    l. Other (please specify)

[if (b) a career in education]

7. Which of the following BEST describes your current career goal?
    a. Science/Math teacher
    b. Other teacher (please specify)

[if (c) a career in other fields]

8. Which of the following BEST describes your current career goal?
    a. Business administration
    b. Lawyer
    c. English/Language Arts specialist
    d. Packaging
    e. Other teacher (please specify)

Commitments

9. Approximately how many <u>hours per week</u> during this semester do you expect to…

   [Options: 0, 1-5, 6-10, 11-15, 16-20, 21-30, 30+]

   a. work at a job this semester/term?
   b. participate in organized extracurricular activities such as sports, college newspaper, or clubs this semester/term?
   c. spend preparing for all classes this semester (studying, reading, writing, doing homework or lab work, analyzing data, rehearsing, or other academic activities outside of class)?
   d. spend preparing for MTH 133 this semester (studying, reading, writing, doing homework or lab work, analyzing data, rehearsing, or other academic activities outside of class)?

Plans and Projections

10. What grade to you expect in MTH133?
    a. 4.0
    b. 3.5
    c. 3.0
    d. 2.5
    e. 2.0
    f. 1.5
    g. 1.0
    h. 0.0.
11. Do you intend to take another math course after this one?
    a. Yes
    b. No
    c. I don't know yet

12. Is another math course required for your major?
    a. Yes
    b. No
    c. I don't know yet
13. How important is a good grade in this course in influencing your decision whether or not to take another math course?
    a. Not important at all
    b. Unimportant
    c. Slightly unimportant
    d. Slightly important
    e. Important
    f. Very important

High School Experience

14. My mathematics courses in high school have prepared me to [Options: Strongly agree, agree, somewhat agree, somewhat disagree, disagree, strongly disagree]
    a. Complete complex calculations without a calculator
    b. Solve word problems
    c. Factor expressions
    d. Solve equations
    e. Solve inequalities
15. The teacher of my last mathematics course in high school [Options: Strongly agree, agree, somewhat agree, somewhat disagree, disagree, strongly disagree]
    a. Lectured most of the time
    b. Primarily showed us how to get answers to specific questions

c. Frequently had us work in groups
d. Frequently had us solve challenging problems
e. Cared that I was successful in the course

Previous Course Experience

16. What was the last math course you took before this one? (excluding statistics courses)
    a. College Algebra / Trigonometry / Pre-Calculus
    b. Calculus I
    c. Calculus II
    d. Calculus III (Multivariable)
    e. Other (please specify)
17. Where did you take that previous math course?
    a. High school
    b. A community college
    c. MSU
    d. Another University
    e. Other (please specify)
18. How long ago did that previous math course end?
    a. 0-1 months ago
    b. 2-3 months ago
    c. 4-8 months ago
    d. 9-14 months ago
    e. 15+ months ago
19. What grade did you receive in that previous math course?
    a. 4.0 — A
    b. 3.5 — A- or B+

c. 3.0 — B
d. 2.5 — B- or C+
e. 2.0 — C
f. 1.5 — C- or D+
g. 1.0 — D
h. 0.0 — E or F

Calculator

20. Please rate the following statements: [Options: Strongly agree, agree, somewhat agree, somewhat disagree, disagree, strongly disagree]
    a. I am comfortable in using a graphing calculator
    b. I am comfortable in using a computer algebra system (e.g., Maple, MATLAB)
    c. I am comfortable with programming (e.g., Python, C++, Java, etc.)
21. In high school I was allowed to use graphic calculators on exams
    a. Always
    b. Sometimes
    c. Never
22. In high school I was allowed to use calculators that performed symbolic operations on exams (e.g, TI-89, TI-92)
    a. Always
    b. Sometimes
    c. Never

Point of View

23. Please rate the following statements: [Options: Strongly agree, agree, somewhat agree, somewhat disagree, disagree, strongly disagree]

a. I believe I have the knowledge and abilities to succeed in this course
b. I under stand the mathematics that I have studied
c. I am confident in my mathematics abilities
d. I enjoy doing mathematics

24. When experience a difficult in my math class… [Scale 1-4]
    I try hard to figure it out on my own ←→ I quickly seek help of give up trying
25. For me, making unsuccessful attempts when solving a mathematics problem is… [Scale 1-4]
    A natural part of solving the problem ←→ an indication of my weakness in mathematics
26. My success in mathematics PRIMARILY relies on my ability to… [Scale 1-4]
    Solve specific kinds of problems ←→ make connections and form logical arguments
27. My score on my mathematics exam is a measure of how well… [Scale 1-4]
    I understand the covered material ←→ I can do things the way the teacher wants
28. If I had a choice… [Scale 1-4]
    I would never take another mathematics course ←→ I would continue to take mathematics
29. When studying mathematics in a textbook or in course materials, I tend to… [Scale 1-4]
    Memorize it the way it is presented ←→ make sense of the material, so that I understand it

30. When solving mathematics problems, graphing calculators or computers help me to… [Scale 1-4]
    Understand underlying mathematics ideas ←→ find answers to problems
31. The primary role of a mathematics instructor is to… [Scale 1-4]
    Work problems so students know how to do them ←→ help students learn to reason through problems on their own
32. Please rate the following statements [Options: Strongly agree, agree, somewhat agree, neither agree nor disagree, somewhat disagree, disagree, strongly disagree]
    a. Mathematics instructors should show students how mathematics is relevant
    b. If I am unable to solve a problem within a few minutes, it is an indication of my weakness in mathematics
    c. In order to succeed in calculus at a college or university, I must have taken it before
    d. Mathematics is about getting exact answers to specific problems
    e. The process of solving a problem that involves mathematical reasoning is a satisfying experience.

About the Labs [post-course survey only]

33. Which of the following did you experience during labs? [Options: Never, Rarely, Sometimes, Often, Always]
    a. You solved a problem using pen/pencil and paper.
    b. You solved a problem by writing or editing MATLAB code.
    c. You analyzed and interpreted results/data to solve a problem.
34. Please indicate the degree to which you either agree or disagree with the following statements about the <u>design</u> of the Application Labs: [options:

(1) strongly agree, (2) agree, (3) neither agree nor disagree, (4) disagree, (5) strongly disagree, or (6) not applicable.]

a. The Labs is a useful tool for learning calculus overall.
b. The Labs help you prepare for quizzes and exams.
c. The Labs help you understand connections between calculus concepts.
d. The Labs help you understand applications of calculus ideas.
e. The Labs help you practice modelling.
f. The Labs make learning calculus seem important.
g. The Labs are important for your success in MTH 133
h. The Labs are important for your achieving your long-term goals at MSU.
i. Weekly quizzes would be more helpful than the Labs for your learning.
j. The length of the Labs is appropriate.
k. Labs should be shorter.
l. Labs should be longer.
m. The difficulty of the Labs is appropriate.
n. The Labs are too easy.
o. The Labs are too difficult.
p. It is easy to determine what the Labs questions are asking.
q. The feedback on the Labs helps you learn.
r. The Labs should be worth a larger part of the course grade
s. The Labs should be worth a smaller part of the course grade
t. The Labs would still be useful if they were not graded

35. Please indicate the degree to which you either agree or disagree with the following statements about working in groups on the Labs: [options: (1)

strongly agree, (2) agree, (3) neither agree nor disagree, (4) disagree, (5) strongly disagree, or (6) not applicable.]

   a. Working in groups on the Labs is enjoyable
   b. Working in groups on the Labs helps you learn
   c. Working in groups on the Labs has helped you make connections for studying outside of class
   d. Working in groups on the Labs is similar to groupwork experiences in the past
   e. Working in groups on the Labs is better than groupwork experiences in the past
   f. Working in groups on the Labs is worse than groupwork experiences in the past
   g. I prefer being assigned to groups
   h. I prefer choosing my own group
   i. Changing groups periodically is a nice way to work with new people
   j. Changing groups periodically disrupts productive group work dynamics
   k. Changing groups periodically alleviates the stress of being in an unproductive group

36. Rate your usual participation in the Labs? [scale 1-4]
   a. I am often very engaged during the Labs ←→ I am rarely engaged during the Labs
   b. I often take the lead during Labs ←→ I rarely take the lead during Labs
   c. I often feel included during Labs ←→ I rarely feel included during Labs

37. Please indicate the degree to which you either agree or disagree with the following statements about MATLAB: [options: (1) strongly agree, (2) agree, (3) neither agree nor disagree, (4) disagree, (5) strongly disagree, or (6) not applicable.]
    a. You can write original MATLAB code to solve problems.
    b. You can modify existing MATLAB code to solve problems.
    c. You can read and understand MATLAB code.

APPENDIX D: LAB DETAILS

1 Lab deployment

Students complete labs in groups of three to four students, randomly assigned at the start of each recitation. Labs consist of a main lab document in the form of a MATLAB LiveScript file, which are e-mailed to students and made available for download on the course website prior to the recitation, and a printed worksheet distributed by the recitation instructor. The MATLAB LiveScript format provides an interactive programming notebook interspersing prose documentation with executable code. Students run the provided LiveScript file either on a local MATLAB installation on their personal laptops, or through the MATLAB Online cloud computing service. Our university has an institutional MATLAB license covering use by all students and faculty. Students follow the instructions given in the MATLAB LiveScript file, and record their answers to factual and interpretive questions on the printed worksheet, which are then collected and graded by the recitation instructor.

The MATLAB LiveScripts and printed worksheets can be downloaded from https://calc2labs.qnlw.info.

2 Application foci of individual labs

For year 3, seven labs were designed with different instructional and application focus.

Lab 1: Numerical Integration

Main learning objective: sums can be used to approximate integrals. This learning objective ties into the definition of the Riemann integral as area under the curve, and provides students with an introduction to numerical integration. Additionally this ties into the “integral test” concept in the portion of Calculus II concerning “sequences and series”.

Lab 2: Baseball Rocketry

Main learning objective: integrals can be used to approximate sums. This learning objective gives the converse of the previous one, showing that in practical applications integrals and sums are mostly interchangeable.

Application setting: derive the relationship between fuel consumed and change of velocity in rocketry (discovering the Tsiolkovsky equation); evaluate which of two two-stage rocket designs has better performance.

Lab 3: Zombie Attacks

Main learning objective: exponential growth and decay, and introduction to differential/difference equations as modeling positive/negative feedback loops.

Application setting: disease transmission in an insular community.

Lab 4: Trial and Tribble-ation

Main learning objective: sequence/series convergence via comparison against geometric sequence/series. Underlying every series convergence test is a comparison against a well-chosen model limiting behavior; this lab ties in with the derivation of the ratio test.

Application setting: algorithm design (run-time guarantees), binary search, signal quantization with error.

Lab 5: Boom Bust Butterfly

Main learning objective: modes of sequence divergence. Address common misconception that for a sequence to diverge the terms must escape to infinity.

Application setting: regular (convergent and periodic) versus chaotic behaviors in market economics forecasting.

Lab 6: Leibniz's Wheel

Main learning objective: the connection between rates of convergence and the radius of convergence of a power series.

Application setting: design a method to efficiently evaluate numerically the natural log function using only addition, subtraction, multiplication, and division; basics of computer science.

Lab 7: Etch-a-sketch Time

Main learning objective: parametric curves, their derivatives, and reconstructions of curves from derivatives.

Application setting: reconstruction of racetrack layout from on-board velocity data on a racecar.

3 Sample worksheet questions for Lab 5

Exercise 1. The number of industry players in the next quarter can be computed from the number in the previous quarter by the formula X(t+1) = R(X(t)) * X(t). The rate of growth function R(X) typically exhibits the following two features: (i) R(X) is a decreasing function of X (ii) R(X) is never less than zero.

Explain why, in the context of our model of a market economy, these two features should be intuitively expected.

Explain also why the solution to R(X) = 1 is considered the market equilibrium.

Exercise 2. [NB: Lab introduced a volatility parameter v into the rate of growth function R(X).]

For X(0) = 0.4, computer the percent change from X(0) to X(1) for the three volatility levels v = 0.5, 1, and 2.

Do the same for X(0) = 2.

You may use MATLAB to perform these computations.

Exercise 3. In the low volatility setting, does the convergence to market equilibrium happen faster for less volatile markets or more volatile ones? How does this compare with our intuitive understanding of volatility?

Exercise 4. In the medium volatility setting, does the convergence to market equilibrium happen faster for less volatile markets or more volatile ones? How does this compare with our intuitive understanding of volatility?

Exercise 5. Run the simulation in the "Chaos" section a few times, with different values of volatility. What do you notice that is different qualitatively between the v = 2.8 case versus the other three cases? What does this say about attempts to forecast stock prices, in a very highly volatile market, based on incomplete information?