Journal of Effective Teaching in Higher Education

Agile and Active: Sustaining Pedagogical Change in a Large-Enrollment Calculus Course

Cynthia A. Cogswell, PhD, Ohio University, cogswell@ohio.edu
Scott Pauls, PhD, Dartmouth College
Adrienne Gauthier, Dartmouth College
Erin DeSilva, Dartmouth College

Abstract. It is well documented that the use of active learning strategies increases student learning (Freeman et al., 2014; Prince, 2004; Springer, Stanne, & Donovan, 1999). A key difficulty in innovating college mathematics is identifying and sustaining what works for both students and the faculty. This study discusses efforts to innovate and sustain curricular change in introductory calculus at a private, elite institution. To examine if incorporating active learning strategies made a difference in student performance, student grades in the redesigned course and performance in subsequent courses were analyzed. Using Austin's 2011 framework to understand the context in which the course redesign took place, individual faculty and contextual barriers and "levers" to sustain change are discussed. Findings are applicable to other STEM disciplines and to colleges and universities in general. Next steps in this research include identifying how to scale change, including, perhaps, networks of faculty to implement and spread the reform on campus.

Keywords: STEM retention, SoTL, Active Learning

Introductory courses play a critical role in introducing students to the content of the discipline and the potential of the field. It is well documented that the use of active learning strategies increases student learning (Freeman et al., 2014; Prince, 2004; Springer, Stanne, & Donovan, 1999), retention of underrepresented populations in the course, and indirectly affects student departure decisions (Braxton et al., 2008; Braxton, Milem, and Sullivan, 2000). Students report greater engagement in courses where instructors signal openness to student questions, discussion, and peer collaboration (Gasiewski et al., 2012; Mazur, 2009). Peer instruction and group work, especially when structured to increase diversity, lead to lower dropout rates among women and racial and ethnic minorities (Blickenstaff, 2005; Cole, 2007; Panitz, 1999; Toppings, 2005; Watkins & Mazur, 2013). Drawing on individual students' experiences and knowledge in class discussions and providing opportunities to challenge the professor's ideas better engage a diverse student body (hooks, 1994).

In a study of over 2000 classes, taught by more than 500 STEM faculty, at 25 institutions, Stains et al. (2018) found that although the methods of and practices in lecturing vary, the lecture is still a very prominent method of instruction. Of the observations, 55% featured lecture 80% of the time. Organizations are looking for ways to create and sustain broad reform, but not much is currently generalizable for how to sustain curricular change once the redesign is complete (Association of Public & Land-Grant Universities, n.d.; Bressoud, Mesa, & Rasmussen, 2015). Innovating college mathematics is not novel; the difficulty is identifying and sustaining what works for each campus context, students, and faculty. Common efforts to integrate active learning into college mathematics have included flipping courses, using adaptive technology, utilizing active learning spaces, and integrating peer learning (Association of Public & Land-Grant Universities, n.d.; Bressoud, Mesa, & Rasmussen, 2015; Fain, 2015; Najmabadi, 2017).

The Gateway Initiative

Dartmouth College is a mid-sized, private, liberal arts institution. A member of the Ivy League, Dartmouth is known to emphasize teaching through a "scholar-teacher" model. In 2014, Dartmouth College undertook its own program to support curricular change. The Dartmouth Center for the Advancement of Learning (DCAL) unveiled an incentive program, named Gateway, that would provide resources to faculty to redesign existing introductory courses. Gateway is designed to address several scarcities in teaching: a flexible and robust budget for teaching assistants, media services, and technology; partnership opportunities with non-faculty educators providing instructional design and project management support; evaluation and assessment services; and finally, status, reward and acknowledgement.

All faculty were invited to apply to Gateway, and among courses selected in 2014 was Mathematics 3, an introductory calculus course. The course annually enrolls ~300 students, with multiple sections offered in both fall and winter terms. Mathematics 3 is a prerequisite to courses in the mathematics, physics, economics, chemistry departments, and more. Unlike other Gateway courses, where faculty teach alone, rotate teaching, or teach in partnerships, Mathematics 3 traditionally has a "teaching team" model, with a single course coordinator who oversees the curriculum and 3-4 new instructors teach sections of the course during the year. Has incorporating active learning strategies made a difference in student performance? What is the context in which this change occurred, and what were the barriers and levers for sustaining change? This study explores these research questions for Mathematics 3.

Mathematics 3

The course coordinator, Pauls, submitted the Mathematics 3 application for Gateway. Once selected as a Gateway awardee, Pauls worked with an instructional designer to revise the course. The revised course uses a mixture of online and written homework sets, where the online platform provides real-time feedback to instructors on student progress. The feedback is aggregated at multiple levels, allowing instructors to adapt class content appropriately. Students also completed problem sets in class in fixed groups of 4-6 members. Over the instances of the course, Learning Fellows (LF) worked in the course to mentor the groups, and each LF was responsible for 3 or 4 groups (Dartmouth Learning Fellows, n.d.). LF's are paid undergraduate teaching assistants recruited to facilitate group work and

interaction. The LF model is based on the work of McHenry, Martin, Castaldo and Ziegenfuss (2010).
Additionally, graduate teaching assistants run drop-in problem sessions and tutorials. This multi-level and multi-modal instructional approach aims at acknowledging and mitigating student heterogeneity upon entry to the course. This structure echoes a principle elucidated by the Dana Center at University of Texas, Austin: "The aim isn't to water down math requirements but to provide 'the right math for the right student at the right time'" (Najmabadi, 2017).

Students at this institution are unusually well prepared in mathematics. The average SAT math score is 727 out of 800 for the class of 2020, and typically over half the incoming class has placement out of one or more courses in the calculus sequences via Advanced Placement, A-level, International Baccalaureate, or local placement exams. Historically, enrollees are a mixture of students who have generally solid pre-calculus preparation, most of whom have seen some or all of the material in the course before, and students with some deficiencies in their pre-calculus training. This heterogeneity presents challenges in aim and scope for Mathematics 3, as well as for placement. The course redesign sought to address these difficulties.

As sections of Mathematics 3 are often taught by a mixture of permanent and temporary faculty, as well as graduate students, the redesign focused significant effort on creating a course infrastructure which minimizes startup time for new instructors. It also provides a degree of uniformity across instances of the course. Since the initial course revision in 2014, all sections of Mathematics 3 share the same syllabus, learning outcomes, a core set of homework sets that are common between the sections, common exams, and final grades that are set by the same rubric. Mathematics 3 sections use small groups and learning fellows (LFs) to facilitate in class discussion and problem solving. Further, faculty designed modular curricular with a menu of active learning components to complement content delivery. Subsequent to each instance, instructors revised and added to content reflecting course experiences. Norms for the course as well as revisions are managed by one course coordinator teaching the course regularly throughout the evaluation period, providing continuity as well as training and guidance for instructors new to the course.

Now, instructors, instructional designers, and LFs have weekly meetings to assess student progress, resolve problems, and engage in training around issues arising in the active learning components of the course. This "just-in-time" component augments baseline training by providing instructors and LFs with practical, timely tools to respond directly to course challenges. These efforts in sameness across section and instructor are aimed at creating environments where students from all sections finish the course with similar mastery of the content and preparedness for what they do next, whether that is in mathematics or some other field.

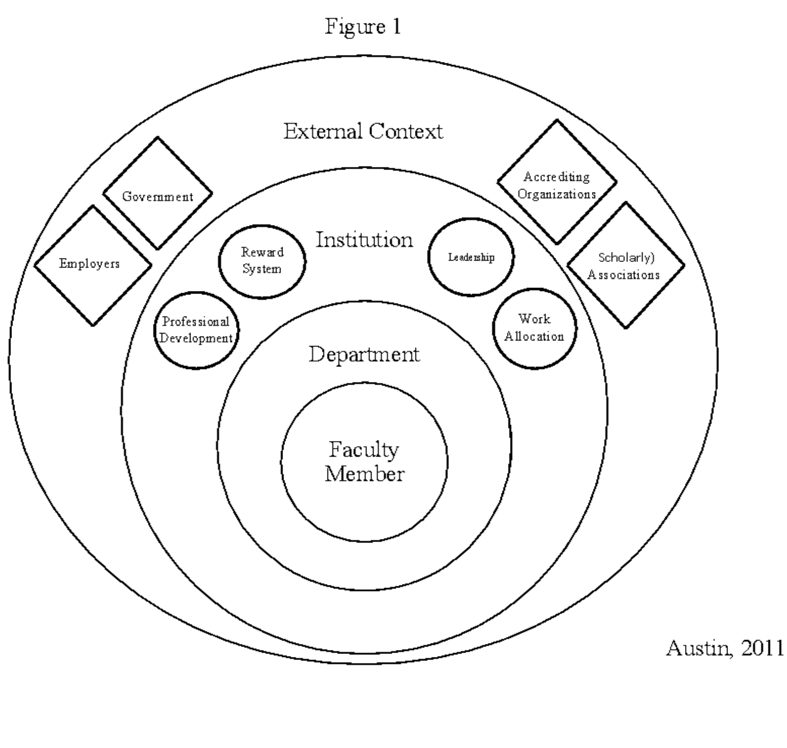
Theoretical/Conceptual Framework

This research uses Austin's (2011) model for understanding faculty members' teaching-related decisions (Figure 1). We use this model to analyze faculty work and choices made regarding teaching. This model is a natural choice as it places strong emphasis on the context, it also has the potential to shed insight, tease out unique characteristics, and identify barriers or levers that have promoted change in faculty behavior.

Methodology

This research took an intrinsic case study approach. In this, the researcher chooses what is to be studied, emphasizing that the case is of primary interest. Stake (2003) wrote, "I call a study an intrinsic case study if it is undertaken because, first and last, the researcher wants better understanding of this particular case." The case is not selected because it represents all other cases or because it emphasizes a particular problem, but because of "its particularity and ordinariness" (p. 136). Case study research seeks to understand human phenomena in all of its bounded complexity. The Mathematics 3 course in the Gateway context is of particular interest because of the way change occurred and has been sustained over three years. This research seeks to understand the particularity and ordinariness of the Mathematics 3 course in the Gateway context.

Figure 1 Austin's Original Model



Case study lends itself to reflective research, multiple data sources, and teasing out the context in which the case exist. This study uses interview data, course artifacts, student grades, and instructor and instructional designer reflections. These artifacts informed the case as they materialized as data sources. Course artifacts, focus groups, and interviews became available at essentially the same time; course grades followed much later; followed even later by this co-constructed paper, making meaning of the case together: the outside-researcher and the inside-participants. Intrinsic case study enables the researcher to interact with the case, and to converse with key players about their lived experiences; which is precisely how this paper was composed.

Data

Data collection began after obtaining IRB approval for this study. To understand the impact of the redesign, student focus groups were conducted (three one-hour focus groups) and faculty were interviewed (three one-hour interviews with the course coordinator). The focus groups asked students about teaching and interactions in class, preparations for class, giving and receiving feedback, and quality and frequency of assessments. Faculty interviews included questions on Gateway, course change, and teaching. We utilized institutional student data from Mathematics 3 between 2009 and 2016, which reflects five years of data before the revision and three years of data with the revision. Additionally, narratives from the course coordinator and instructional designers were collected. The research questions are examined separately followed by a discussion.

Measuring Impact: Student Performance

To answer our first research question (Has incorporating active learning strategies made a difference in student performance?), we compared student performance in the revised course to a cohort of students from a five-year window before the revision, utilizing descriptive statistics as well as two-sample t- and Kolmogorov-Smirnoff (KS) tests. The KS-test is less well known than the t-test; it compares two samples and evaluates whether they have been drawn from the same unknown distribution (Corder & Foreman, 2014). Overall grade distributions differed slightly. Quartiles for the revised course were (C, B-, B+) while those for the five-year comparison window were (C, B, B+) but the distributions cannot be distinguished statistically by either the t- or KS-test. We see a rise in the number of withdrawals - students who choose to drop the course with penalty - in the revision where 8.3% of students withdraw versus 5.8% in the historical instances. Withdrawal rates give us our initial measure of persistence as it shows the percentage of students who exit the course before completion. Table 1 shows withdrawal rates across several demographic categories while Table 2 shows the raw counts. While in all but one category - Asian/Asian American/Pacific Islander - withdrawals rose in the revised course, the magnitudes were unevenly distributed. Women withdrew at lower rates than men. Among ethnic categories, Hispanic and Latinx students fared the best in the revised course with only a slight uptick in withdrawal rates. White (non-Hispanic), black (non-Hispanic), and multi-racial students' withdrawal rate increases are almost identical to the overall increase, while Native American students fared worse.

We examined longer term persistence by looking at outcomes of students who took the sequel calculus course in the term after completing the target course. Historically, these persistence rates are 14.9% while the revised course saw a rate of 15.3%. Student grade outcomes in the sequel calculus course to the revised course were significantly better than those of comparable students historically (see Table 3). The two-sample t-test is significant ($p<0.01$) as is the two-sample KS-test ($p<0.05$) for the sequel course grades when restricted to students who got less than an A in the target course. Over the entire range of grades the p-value is approximately 0.06 (t-test) and 0.15 (KS-test).

Table 1

Withdrawal rates for students in revised and unrevised versions of the target course

Course	All	Female	Male	AAAPI	BNH	HL	MR	NA	WNH
Unrevised target	5.8%	4.6%	6.9%	2.3%	7.3%	8.4%	10.0%	16.2%	4.7%
Revised target	8.3%	6.8%	9.8%	1.8%	10.2%	8.9%	12.8%	20.3%	7.4%

Percentages of students withdrawing from the target course, broken down by gender and ethnic/racial categories. AAAPI=Asian/Asian American/Pacific Islander; BNH=Black, non-Hispanic; HL=Hispanic/Latinx; MR=multi-racial; NA=Native American; WNH= White, non-Hispanic.

Table 2

Withdrawal totals for students in revised and unrevised versions of the target course

Course	All	Female	Male	AAAPI	BNH	HL	MR	NA	WNH
Unrevised target	83 (1428)	31 (671)	52 (755)	4 (172)	17 (231)	13 (155)	2 (20)	11 (68)	31 (656)
Revised target	80 (958)	32 (468)	48 (489)	2 (113)	17 (166)	10 (112)	5 (39)	15 (74)	31 (418)

Raw counts of students withdrawing from the target course (total enrollment in parentheses), broken down by gender and ethnic/racial categories. AAAPI=Asian/Asian American/Pacific Islander; BNH=Black, non-Hispanic; HL=Hispanic/Latinx; MR=multi-racial; NA=Native American; WNH= White, non-Hispanic.

Table 3
Grades in a sequel course conditioned on grades in the target course

Letter Grade Target Course	Grade Quartiles sequel course (after unrevised target)	Grade quartiles sequel course (after revised target)
D	(F,D,D)	(F,C-,C)
C-	(F,F,D)	(F,C+,C+)

C	(F,D,C)	(C,C+,C+)
C+	(F,D,C+)	(C-,C+,B)
B-	(D,C,B-)	(C,C+,B)
B	(C,B-,B+)	(B-,B,B+)
B+	(B-,B,B+)	(B-,B+,B+)
A-	(B,B+,A-)	(B,B+,A-)
A	(B+,A,A)	(B+,A,A)

We show the median grade in the sequel for students coming from the target with a specific grade. We see statistically significantly higher grades in sequel for students with lower grades (less than A) in the revised target than similar students from the unrevised version (two sample t-test, p<0.01, two sample KS-test, p<0.05)

Analyzing the Case

In this section we use Austin's (2011) framework to examine the context in which this case took place, addressing the second research question, what is the context in which this change occurred, and what were the barriers and levers for sustaining change? Austin's model, Figure 1, involves labeled concentric circles, layering individual background characteristics, values, and training, within the organizational environment and all its many potential barriers or 'levers' to promote change in faculty behavior. The analysis is organized in three sections: (1) relevant individual characteristics, (2) contexts affecting faculty practice, and (3) barriers and/or levers impacting faculty teaching practices.

Intrinsic case study lends itself to reflective research, multiple data sources, and teasing out the context in which the case exist. The following sections do just this. To examine the layers within Austin's framework, Cogswell and Pauls wrote the analysis of the case together. Cogswell drafted the initial analysis, Pauls, reviewed, edited and provided feedback. After multiple exchanges and conversations, they arrived at a shared description of the case. The paragraphs below reflect their work; Pauls drawing on his experience and Cogswell drawing on interview data, course artifacts, student focus groups, and end of term student evaluations.

Relevant Individual Characteristics

Prior Experience. Pauls is a tenured faculty member. He earned his PhD in Mathematics from University of Pennsylvania and bachelors in Mathematics from Columbia University. During his graduate coursework he received recognition for teaching several times over (5 terms). At the outset of the redesign, Pauls had over 15 years of teaching experience in mathematics. Calculus is a core mathematics course - the "bread and butter" of most departments as it draws the highest enrollments. As with most mathematicians, he has taught calculus at several levels regularly throughout his career.

Prior to Gateway, Pauls taught Math 3 and all of his courses using predominately lecture and or interrupted lecture formats, and the latter only in the last few years before the Gateway redesign as he began to experiment with different approaches. When applying to be a part of Gateway, he mentioned wanting to integrate video instruction into orientation and placement materials for the course. The proposal Pauls submitted details that student and faculty evaluation of the course indicate that "the traditional lecture format hinders progress of a substantial subset of students," which he stated, "contributes to both frustration and poor learning outcomes."

At Dartmouth he has served as vice chair and then chair of the Department of Mathematics. In addition, he has served on a number of campus committees, including a first-generation enrichment program steering committee, advisor to first year students for mathematics, mathematics recruiting committee chair, and more. Pauls seems to prioritize teaching and improving access to the Mathematics field for all individuals.

Doctoral Socialization. Graduate programs oriented towards research careers do not typically have extensive pedagogical training. While programs view competence in the classroom as necessary for successful students, they do not necessarily invest resources in this direction. Pauls' pedagogical training consisted of a series of seminars where advanced graduate students mentored newer students and introduced them to teaching at the institution. The last seminar, shortly before classes started, included several faculty members who brought their perspectives. Methodologically, the seminars presented only lecture oriented teaching. Consequently, Pauls' teaching was oriented solely towards lecture for the majority of his career. However, noteworthy from Pauls' doctoral socialization is that while teaching in graduate school he received the Moez Alimohamed Graduate Student teaching award, as well as four departmental teaching awards.

Discipline. Austin (2011) wrote, "Disciplines have distinct cultures, including values and criteria about what constitutes excellent work and norms for the behavior of members of the field (Austin, 1994, 1996; Becher, 1987)" (p. 7). The Mathematics discipline at Dartmouth College approaches teaching in alignment with the field broadly. It neither encourages nor dissuades attention from teaching. Instructors and faculty have conversations about teaching, just like they do about their research.

Career Stage and Appointment Type. Austin's framework states that "early career faculty members and doctoral students show that they are often eager to share their passion for their discipline and fields with novice learners," but Pauls is not an early career faculty member. Pauls is a tenured full professor with many years of classroom experience. With tenure and promotion in the past, Pauls has the flexibility and experience to devote time and energy to this curricular project. Yet, Pauls still has enthusiasm to "share [his] passion for [the] discipline and fields with novice learners," as evidenced in his voluntarily undertaking this project. Pauls shared that he took on this project to improve the curriculum of the course and to create a better alignment between student readiness and course content.

Pauls believes that if the mathematical community wishes to bring more people into mathematics (and STEM more broadly) and address the problems of uneven representation, then we have no choice but to rethink how we educate and train students. Such solutions will not be quick or easy, but rest on sustained effort, evaluation, and revision.

Faculty Motivation. An historical look at student outcomes prompted the beginning of this work. Pauls' started the work on improving Mathematics 3 before he knew of the Gateway Initiative. Analyzing student outcomes both formally and anecdotally showed the unsurprising result that calculus instruction, including his own, was not terribly effective. Particularly troubling were the results for students from groups underrepresented in mathematics, who disproportionately left the field. These analyses prompted the revision.

Moreover, other efforts to incorporate new pedagogical methods - particularly those under the umbrella of active learning - at Dartmouth provided examples and templates for work in mathematics. Support from DCAL multiplied the benefits of the examples by providing support and research tools. To some degree, one could question how much the support from the Gateway Initiative influenced what happened. If the Gateway Initiative did not exist, would as much work as has been done taken place? Probably not.

Contexts Affecting Faculty Practice

In this section the next layers of Austin's framework, the institutional, department and external contexts, are discussed.

Institutional Context. As stated earlier, Dartmouth College is a mid-sized, private, liberal arts institution. A member of the Ivy League, Dartmouth is nearly 250 years old. Sometimes the age of the institution is used in jest to explain why change is slow on campus. As Pauls stated, scholarship, teaching, and service are considered in tenure review in the order listed. How teaching is reviewed varies but it is through a combination of a self-authored essay, seeking comments from past students, and a review of end of term student evaluations of teaching.

Beyond tenure review, excellent teaching is recognized on campus formally and informally. Formally, the Dean of Faculty annually recognizes about ten faculty for instruction and there are five named teaching awards. The awardees are selected by either the Dean of Faculty, Deans of the College, or the Dartmouth Center for the Advancement of Learning. These awards do not include significant reward (e.g. a course buy-out), but some include a stipend.

Informally, faculty learn of excellent teaching through conversation. They learn from conversations amongst themselves, within and across departments. Campus news and blog posts feature teaching innovations and those who have received awards. Also, faculty learn of and about teaching from conversations and events with instructional designers.

Departmental Context. Mathematics faculty value strong and effective pedagogy. Dartmouth faculty reflect this commitment, aiming to excel both in their research and pedagogical endeavors. This context provided both motivation and support for the curricular and pedagogical revision.

Many faculty experiment with novel teaching approaches to fine tune their instruction and, by mid-career, typically settle upon a collection of techniques they find most effective for the course they teach. The Department Chair assigns courses each year based on faculty requests balanced across curricular needs. The Chair also assigns experienced permanent faculty as course supervisors for courses where less experienced instructors are teaching. Junior faculty have teaching mentors (both formally and informally) who visit classes, review materials, and give advice. Course supervisors mentor new post-doctoral instructors and other visitors as they join the department and the teaching faculty.

Dartmouth College's graduate program is something of an exception, where students are required to work through a rigorous theoretical and practical pedagogical curriculum before entering the classroom at teachers. Many introductory courses use advanced undergraduates as graders and often have graduate students assigned as teaching assistants. The Learning Fellows program extends pedagogical support teams which, in particular, makes it feasible to run intensive active learning in larger classes.

Creating the new structure for Calculus and refining the active learning components to fit Dartmouth College's students and curriculum does present a barrier for broader adoption. Some members of the Mathematics faculty are still skeptical of using these approaches in their own classrooms. Two factors have helped in broadening interest: our analysis of outcomes and demonstration of success in later parts of the curriculum, and our construction of a library of materials and methods for bringing active learning into the classroom. The first helps convince skeptics of the usefulness of the methods, particularly in the context of courses in their own department while the second lowers the time barrier for adoption. The latter has also been very helpful when bringing in new and/or less experienced instructors to the class. Of course, this has meant more work for Pauls and his team.

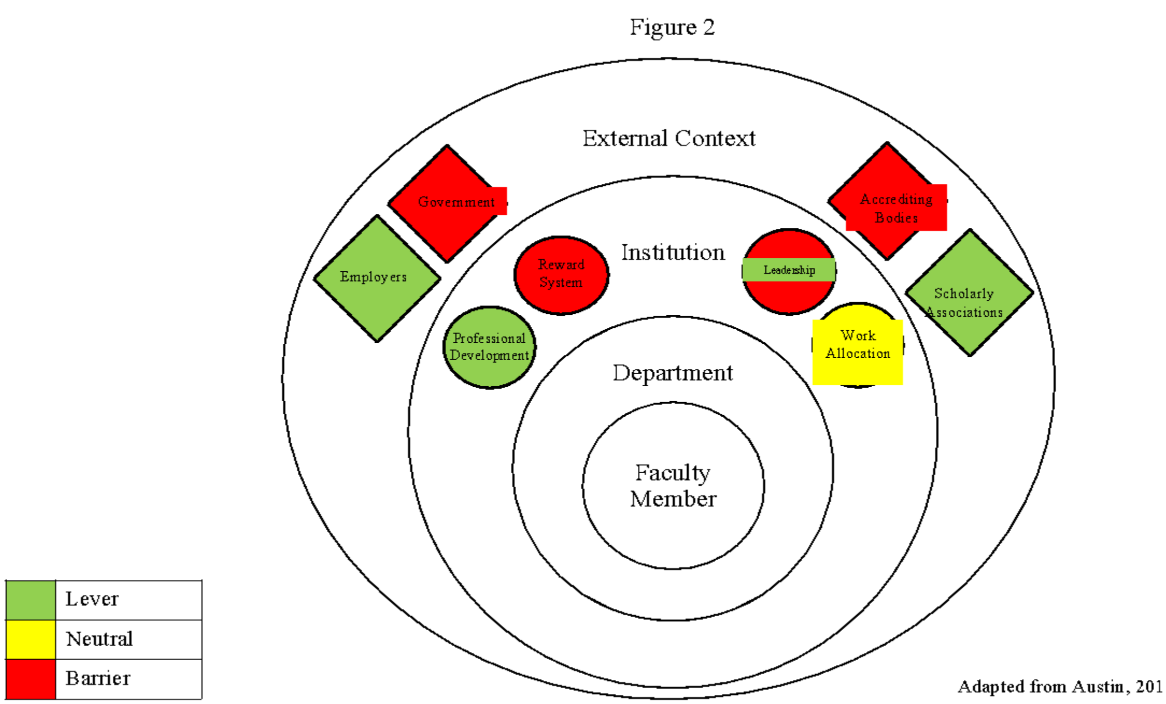
External Contexts. Within the external contexts of government, the federal government leaves much of the review of educational quality to accreditors. Dartmouth College is accredited and in good standing with its regional accreditor, the New England Commission of Higher Education (NECHE) (2019). Among accreditors, NECHE is known for allowing institutions more flexible expectations for student learning outcomes and measures. In compliance with NECHE, Dartmouth degree programs have learning outcomes posted online. The mathematics department is no exception, and has degree outcomes and expectations clearly articulated. The oversight from the state government is minimal, and mostly directed at initial review of institutions seeking establishment, as opposed to ongoing review (New Hampshire Department of Education, 2012). For these reasons, both the accreditor and the government are barriers to change.

On behalf of the Mathematical Association of America, Bressoud, Mesa and Rasmussen (2015) edited a volume of recommendations for content and instruction of college calculus. Foremost, their work calls attention to the need to make calculus more accessible to all. In addition, Saxe and Braddy (2015) looked at seven curricular guides published by five mathematical professional associations to reconcile their recommendations and outline a course or reform, whose fundamental finding for mathematics curriculum and instruction was that “the status quo is unacceptable” (p. 1). Additional authors have had similar recommendations (Association of Public & Land-Grant Universities, n.d.; Fain, 2015; Najmabadi, 2017). From this review and the existing research and support from these groups, we determine the scholarly associations to be levers for change. The last external context in Austin’s model are employers. The National Task Force (2012) scanned the literature, held a series of roundtables, and vetted their findings with employers on what they desire from collegiate graduates. They concluded with five recommendations, all demanding and valuing greater emphasis on a range of student learning outcomes and competencies. For this reason, we concluded that employers are levers for change.

Barriers and/or Levers Impacting Faculty Teaching Practices

The next step in applying Austin’s theory to this case is to identify if the rewards system, work allocation, professional development and leadership are barriers or levers to change. While some of these have already been described, in the paragraphs that follow they will be discussed and labeled as either barriers to innovation or levers facilitating innovation. Please refer to Figure 2 as a guide to reading the following paragraphs.

Figure 2
Austin’s Model in the Mathematics 3 Case



Reward Systems. The rewards systems in place at Dartmouth College include: tenure and promotion, teaching awards, grants to fund teaching innovations, and informal recognition. Austin (2011) writes, "higher education institutions are sending strong messages about the relative value of time spent on research versus time spent on teaching" (p. 11). There is a lot of pressure on faculty to do research. There are not published, clear guidelines or expectations for scholarship, teaching and service in the tenure and promotion process, thus clouding what the emphasis should be on each part, and leaving it to each individual to determine how much time to allocate to each effort.

How faculty respond to these institutional signals and messages varies. Broadly, tenured faculty mentor tenure track faculty, and tend to give them less burdensome courses so that they can devote more time to research and publications. Teaching awards, discussed earlier, do not have significant resources attached to them and they do not carry enough provenance to be a reason in of themselves to teach well. As stated earlier, while Dartmouth expects a high level of pedagogical excellence, the tenure and promotion system does not set research, teaching, and service on equal footing. One could reason that Pauls was able to execute this revision only after tenure - work on such a project earlier would likely have hurt the tenure decision.

In short, the rewards systems in place, purposeful or not, are barriers to curricular change. The lack of clarity on tenure and review fails to promote the importance of teaching and does not ascribe clear expectations, or values, for what is considered bad, good, or exceptional teaching.

Work Allocation. Faculty teaching expectations are decided on a departmental basis. Faculty within a single department have uniform teaching loads set by the dean of their division, mitigated only by course buy-out. The number of courses taught per year by individual is 5, 4 or 3. Broadly, the expectations are the same for tenure-track and tenured faculty.

In contrast to the national context, Dartmouth has very few adjunct faculty and rarely permits graduate students to teach. In mathematics, advanced graduate students do teach sections of Mathematics 3 but with required training, and now because of Gateway, with out-of-class and in-class support. Additionally, Dartmouth is on the quarter system so faculty teaching loads are divided across three terms, leaving faculty with the fourth term "off," which is commonly devoted to research. For these reasons work allocation is at best a secondary condition for change. If teaching loads were to be reduced, systematically or a one-time recognition of work for innovative teaching, one could easily reason that the extra time would go to research.

Professional Development. Faculty have access to professional development for their teaching. On campus, the Dartmouth Center for the Advancement of Learning (DCAL), the library, and educational technologies offer teaching resources. Faculty can sign up for a one-time workshop on a topic, join a reoccurring book club, and speak one-on-one about teaching related topics. DCAL also offers modest funding for faculty to present on teaching at conferences away from campus. Recently, DCAL had an influx of funding which has been distributed through the experiential learning initiative and Learning Fellows (LF) program— efforts that followed after Gateway. Similar to Gateway, faculty and staff can receive funding to support teaching innovations, and to have LFs in their classroom. These professional development opportunities have offered a faculty community, team approaches to teaching, and access to funds supporting instruction (e.g. to record videos).

Additional support for faculty development are the instructional designers on campus. As stated earlier, instructional designers work with faculty, supporting them in any pedagogical changes they wish to undertake, from experimenting with one assignment to an entire course redesign. The instructional designers are levers to change.

Another teaching opportunity, the Dean of Faculty offers funding for two faculty from differing departments to create a cross-division course. If faculty want to co-teach or team teach a course, they submit a course proposal to the Dean of Faculty. This initiative is one of the only ways to co-teach a course while still receiving full teaching credit. Related, if a faculty member wants to create a new course, the course must first be reviewed and approved by its aligning department or program, then the registrar, then be reviewed by the Divisional Council, and finally, reviewed and approved by the faculty Council on Instruction.
From our analysis, professional development is a lever to change but only if the faculty member is interested in it themselves. It is a small lever. There are resources available if a faculty member seeks them out and values interacting with others about their teaching.

Leadership. Austin (2011) said of leadership as a barrier or lever that "leaders at each level of the institution are important in creating a culture that encourages, supports, and rewards teaching innovations that support student learning" (p. 14). She wrote that provosts can signal the importance of teaching through their leadership, and can provide guidelines for tenure and review, "emphasizing the importance of commitment to teaching excellence as part of the review criteria and the expectations for success" (p. 14). Likewise, deans and department chairs influence what is valued and can provide support and incentives for professional development.

Institutional leaders have not clearly defined the role of teaching in tenure review, other than to indicate its importance. The faculty handbook states amidst the expectations for promotion to associate professor, "It is not possible to enumerate specific qualifications for tenure so precisely and objectively that the need for judgement is obliviated" and, "It is difficult to define outstanding teaching in specific terms" (Dartmouth College, 2016, p. 34). However, the language for evaluating scholarship is specific, stating, "The qualitative assessment of books and articles and of artistic and other professional accomplishments or contributions to the larger scholarly community will be more consequential than the quality of the work," defining a hierarchy of values.

Neither have the leaders illuminated the expectations for teaching beyond tenure. Austin argued that a key element that fosters leadership as lever is "the presence of an institutional leader who serves as a champion, is committed to the overall [scholarship of teaching and learning] goals, and has sufficient institutional seniority to allocate institutional resources and time" (p. 14). Deans and chairs do not oppose innovations in teaching, but they have limited (if any) discretionary funding to support the cost of teaching innovation or to reward faculty for exceptional work. However, as a lever, the Dartmouth Center for the Advancement of Learning does have the positioning and the funds to award faculty opportunity to innovate teaching. These funds have come from institutional leadership, so one could argue that there is executive leadership, it is merely diffused.

Dartmouth currently has some incentives to facilitate this type of innovation. The Gateway initiative is one of them - providing Pauls with design and pedagogical support, financial support for Learning Fellows, and a like-minded community of faculty. Further incentives, as mentioned earlier, the number of instructional designers on campus has doubled in recent years. However, this support is peripheral in Austin's "reward levers" – which emphasizes tenure, promotion, and salary.

A second incentive, there is an institutional team that meets regularly and works together effectively. There is a STEM retention committee, which includes STEM faculty and institutional leaders. The group meets regularly to discuss current research, analyzes student retention behaviors, and discusses interventions to better support undergraduate STEM students. Beyond this committee, there is not a clear institutional vision for the goals to be accomplished. Other than a general charge to continue excellent research and teaching, there is not clear direction on where energies should be directed. We view the administrative processes as barriers to change but the current leadership as levers. From our analysis, leadership is both a barrier and a lever.

Discussion

This paper examined the impact of the redesign by examining student grades, followed by an analysis of the context in which the redesign took place. In this section we discuss the fit of Austin's framework to this case study, followed by a discussion of the intervention's impact on student grades and academic performance.

Austin's Model as a Framework for this Case

Austin (2011) challenges the assumption that change resides solely with the faculty member. Earlier research by Fairweather (1996) examined faculty attitudes and "found that faculty perceived their rewards to be dependent on research, not teaching, including faculty from institutions with a strong emphasis historically on teaching" (p. 46). Individual faculty members make teaching decisions through and within their values, backgrounds, abilities, and aspirations. Institutional change often does not happen because of barriers that are insurmountable, lack of incentives, recognition, or reward.

Why did Pauls' contribute so extensively to undergraduate instruction? Pauls has tenure. So, further, why did someone with tenure undertake a project such as this? The individual faculty member, and perhaps also the departmental context, which were not evaluated as levers or barriers, seem to have tipped the scales as course change did take place despite the counting of barriers versus levers.

In this analysis of Mathematics 3, there were more barriers to innovation than levers (see figure 2). There was no single layer of the institution that was entirely a lever or entirely a barrier; these permeated the often "top-down" or "bottom-up" approaches to institutional change. Further, it is up to our interpretation to define to what extent different layers within the model should be weighed.

Some factors encouraged adaptation (e.g. professional development, scholarly associations, and employers), other factors discouraged innovation (e.g. the reward system, accrediting organizations, and the government). While this analysis concluded that there were more barriers to change than levers for change, the framework is not prescriptive. Austin outlined the multiple contexts, and in this analysis we filled in our understanding of the influences and interactions of various systems.

Next steps in this research include identifying how to scale this change, including, perhaps, networks of faculty to implement and spread the reform at Dartmouth College.

Student Intervention and Impact on Grades and Performance

Over the four instances of the course, instructors tested several different combinations of in-class work ranging from almost entirely active learning techniques to an equal balance of lecture and active learning activities. Student feedback - both in midterm and end-of-term surveys - indicate roughly equal preferences for the mixed approach. Student focus groups and anecdotal follow-up shows that students with high levels of competence in calculus topics prior to taking the course hold the second preference. Consequently, we conjecture that preparation and confidence in the material differentiates between the two preferences.

The overall increase in withdrawal rates is potentially evidence of the failure of the revision to decrease persistence but could also be due to other factors, such as failure to adequately advise and place incoming students. However, the impact on different demographic subgroups of students provides evidence that the revision is having an impact on persistence and success. In particular, the revised course looks particularly effective among the groups of Asian/Asian American/Pacific Islander and Hispanic and Latinx students. Moreover women, black (non-Hispanic), and multi-racial students seem to benefit from the revision more than others. We consider these results suggestive but not yet compelling. As advising and placement align more completely with the revised course, we expect future instances of this course to more fully answer the question of student persistence and the effectiveness of the methods utilized in the revision.

Student grades in the sequel calculus course compared to historical instances provide the most compelling evidence that the course revision provides students with better tools for learning and retaining mathematics. Students who receive lower grades in the target course have increased performance in the sequel course compared to the historical cohort, suggesting that the success of the redesigned course lies in overcoming under-preparedness in students.

Given the heterogeneity of student preparation in the course, instructors found that an approach balancing active learning with lecture in roughly equal proportions proved most effective in terms of student satisfaction. Outcomes, as measured by levels of success in sequel courses, are not significantly different between the instances with different delivery modes in the redesigned course, but the lack of significance may be due to small sample sizes.

As results from mid-term questionnaires and end-of-term evaluations are mixed, we still have the opportunity for further gains. A central approach focuses on better initial student placement. At the beginning of the course redesign, the institution offered an alternate, two-term calculus sequence which replaced the target course for students who needed review of algebra and trigonometry alongside an introduction to calculus. After two instances of the redesign, we realized that the heterogeneity of preparation among students in the course was detrimental to overall positive outcomes, reflected in grades and end of term evaluations. Consequently, we redesigned the sequence in the department to have a single course aimed at students who had seen no calculus before. In the following two instances of the redesigned course, student preparation heterogeneity significantly decreased resulting in more uniform course outcomes. We expect continued positive gains from this change as campus-wide advising on mathematics placement permeates the collective student awareness.

A revolving instruction model brings challenges to sustaining curricular change (Jones & Harris, 2012; Society for Industrial and Applied Mathematics, 2012). Consistency issues in curriculum, and what students learn across different sections of the same course, arise when sections are taught by multiple instructors. One of the limitations in college mathematics instructional models is that large sections of entry-level courses are taught by graduate students, postdocs, or visiting professors. It is the same model at this institution, where new hires are asked to teach Mathematics 3 their first term on campus and teach it at most 2 or 3 more times before they leave the institution or switch to another courses. With Pauls' as the course coordinator, the changed curriculum is both sustained and extended to the revolving door of instructors—extending the investment in the course to more instructors and thus more students.

Agile, Active, and Vulnerable

From this examination, it is evident that incorporating active learning strategies made a difference in student performance, and through applying Austin's (2011) framework, a new understanding of the context in which change occurred, and the barriers and levers for sustaining change explored.

Prior to listening to how faculty change, in our experience, audiences want to know if the student learning improved. In this course, it did. Pedagogical innovation is messy and matching learning design to student needs, course outcomes, and preparedness takes time. The Mathematics 3 course described here is of particular interest because change has been sustained over three years. Deslauriers et al. (2019) set out to study if the students and instructors who felt they learned less in active learning environments than their lecture-based peers' feelings were true and found they were not. They discussed how a superb lecturer could create a greater feeling of learning than the more effective active learning. Students feeling that lecture is more informative rails against the push towards active learning, and reinforces the vulnerability of the Mathematics 3 revisions. The Learning Fellows, regular teaching huddles, and continued investment of Pauls' time all support the changes to the course.

Absent strong, stable levers, the pedagogical innovations are vulnerable—easily lost should Pauls leave his role and equally lost should the funding for Learning Fellows or instructional designers be unsustained. While Gateway is likely a temporary blip in inspiring innovation, this research has documented and shed light on why it was effective in order to 'save' the learning for future pedagogical innovation programs or levers.

Institutions are complex organizations. Change is hard, value laden, and not a linear trajectory. In this case Pauls is the singular continuous thread across the narrative and across the changes in Mathematics 3. Cogswell has considerable concern over what would happen should Pauls leave his position. But the purpose of using Austin's framework was to shed light on why the change even happened in the first place. When Pauls began this work, pedagogical innovation at Dartmouth was not commonplace. The addition of the Gateway program added a cohort of others that are engaged in similar work, dissolving some existential angst.

Limitations

Although there are some limitations to this study, including the heterogeneity of the population, and the assumption that grades are an indicator of learning and course revision success, this study contributes to the higher education research on revising mathematics curricula. While the course is still being taught, it is still evolving, and this research provides a case account of course change, impact, and how change has been sustained.

Implications and Future Research

Research on the process and impact of active learning at Ivy League institutions is limited. Through this study we documented how an entry-level course changed its curriculum, sustained it over time, and what impact the changes have had on student performance. This study serves to add a deeper understanding of sustaining course change in introductory mathematics courses by examining the context in which change occurred.

Expanding on this work, implications for future research include direct observation of classes, interviews with other Mathematics 3 instructors, and direct gains measurement of student learning in the course. Additionally, because of the emphasis on small group work in Mathematics 3, examining student confidence in mathematics could be insightful.

Conclusion

Higher education practitioners and scholars have been talking about change and reform in science, technology, engineering and mathematics education for three decades (Kezar, Gehrke, & Elrod, 2015). Often, the topics of tenure and pedagogical change are discussed together (Tagg, 2012). Stains et al. (2018) recommend that institutions revise tenure, promotion, and merit-recognition policies to align with evidence-based instruction. The reward of tenure is the often-blamed crux of tension between research and teaching as it is, in the words of Fairweather (1996), "the very structure which ensures maintenance of the disciplines [and] works against faculty involvement in teaching and learning and against developing a more successful undergraduate curriculum" (p. 105). The department is the preserver of the discipline and individual space, with tenure rewarding contribution to the discipline and teaching as a tax to the institution. This paper contributes to a research void, with few accounts of institutional change efforts (Kezar, Gehrke, & Elrod, 2015). If we're going to consider how to reform STEM education, reform can be sparked by efforts from national, state, institution, departmental contexts, but it will not be sustained if the faculty member is not at the center.

Changing pedagogy and thus student learning in science, technology, engineering and mathematics has been a call to action for decades. Yet the issue persists. We studied Mathematics 3 to better understand the impact curricular change has had and to better understand how and why the change occurred. It is the authors' hope that by documenting this case and using Austin's (2011) framework we contribute to the work and research on teaching interventions.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

References

Austin, A. E. (2011). *Promoting evidence-based change in undergraduate science education.* National Academies National Research Council. Retrieved from tidemarkinstitute.org

Association for the Advancement of Science [AAAS]. (2011). *Vision and change in undergraduate education: A call to action*. Retrieved from http://visionandchange.org/files/2011/03/Revised-Vision-and-Change-Final-Report.pdf

Association of Public & Land-Grant Universities. (n.d.). Student engagement in mathematics through a network for active learning (seminal). *Association of Public & Land-Grant Universities.* Retrieved from http://www.aplu.org/projects-and-initiatives/stem-education/seminal/

Adelman, C. (2015). To imagine a verb: The language and syntax of learning outcomes statements. *National Institute for Learning Outcomes Assessment, February 2015*. Champaign-Urbana, IL: National Institute for Learning Outcomes Assessment.

Blickenstaff, J. C. (2005). Women and science careers: Leaky pipeline or gender filter? *Gender and Education, 17*(4), pp. 369-386.

Braxton, J. M, Jones, W. A., Hirschy, A. S., & Hartley, H. V., III. (2008). The role of active learning in college student persistence. *New Directions for Teaching and Learning, 115*, Fall 2008, pp. 71-83.

Braxton, J., Milem, J., & Sullivan, A. (2000). The influence of active learning on the college student departure process: Toward a revision of Tinto's theory. *The Journal of Higher Education, 71*(5), pp. 569-590. doi:10.2307/2649260

Bressoud, D. (2015). Insights from the MAA national study of college calculus. *Mathematics Teacher, 109*(3). Retrieved from http://www.nctm.org/Publications/Mathematics-Teacher/2015/Vol109/Issue3/Insights-from-the-MAA-National-Study-of-College-Calculus/

Bressoud, D., Mesa, V., & Rasmussen. C. (2015). *Insights and recommendations from the MAA national study of college calculus.* MAA Press. Retrieved from: http://www.maa.org/sites/default/files/pdf/cspcc/InsightsandRecommendations.pdf

Brown, M. B. (2012, May 5). *Class flipping: Good pedagogy or online hype?* (Blog post). Retrieved from http://markbbrown.com/2012/05/05/class-flipping-good-pedagogy-or-online-hype/

Cole, D. (2007). Do interracial interactions matter? An examination of student-faculty contact and intellectual self-concept. *Journal of Higher Education, 78*(3), 248–272.

Corder, G. W. & Foreman, D. I. (2014). *Nonparametric Statistics: A Step-by-Step Approach.* 288pp. Wiley.

Dartmouth College (2016). *Handbook of the faculty of arts & sciences*. Hanover, NH. Retrieved from https://faculty.dartmouth.edu/dean/sites/faculty_dean.prod/files/dean_faculty/dartmouth_fac_handbook.pdf

Dartmouth Learning Fellows (n.d.). *Home.* Retrieved from https://sites.dartmouth.edu/learningfellows/

Deslauriers, L, McCarty, L. S., Miller, K., Callaghan, K. & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom*. Proceedings of the National Academy of Sciences of the United States [PNAS], September 24, 2019, 116*(39), pp. 19251-19257. Retrieved from https://www.pnas.org/content/116/39/19251

Fain, P. (2015 August 28). *Learning to adapt.* Inside Higher Ed. Retrieved from: https://www.insidehighered.com/

Fairweather, J. S. (1996). *Faculty work and public trust: Restoring the value of teaching and public service in American academic life*. Allyn and Bacon: Boston, MA.

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

Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. J. (2012). From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Research Higher Education, 53*, pp. 229-261.

hooks, b. (1994). *Teaching to transgress: Education as the practice of freedom*. New York, NY: Routledge.

Jones, F. & Harris, S. (2012). Benefits and drawbacks of using multiple instructors to teach a single course. *College Teaching, 60*, pp. 132-139. Routledge.

Kezar, A., Gehrke, S., & Elrod, S. (2015). Implicit theories of change as a barrier to change on college campuses: An examination of STEM reform. *The Review of*

Higher Education, 38(4), pp. 479-506. Association for the Study of Higher Education.

Mazur, E. (2009). Farewell, Lecture? *Science, 323*, 50-51. Retrieved from http://science.sciencemag.org/content/323/5910/50

McHenry, N., Martin, A., Castaldo, A., & Ziegenfuss, D. (2010) Learning assistants program: Faculty development for conceptual change. *International Journal of Teaching and Learning in Higher Education, 22*(3), pp. 258-268.

Najmabadi,S. (2017 April 16). Math gets a makeover. *The Chronicle of Higher Education.* Retrieved from: http://www.chronicle.com/

The National Task Force. (2012). *A crucible moment: College learning & democracy's future.* Association of American Colleges and Universities. Washington, D.C. Retrieved from

https://www.aacu.org/sites/default/files/files/crucible/Crucible_508F.pdf

New England Commission of Higher Education. (2019). *Colleges and Universities.* Retrieved from https://www.neasc.org/colleges-and-universities

Panitz, T. (1999). Benefits of Cooperative Learning in Relation to Student Motivation, in Theall, M. (Ed.) Motivation from within: Approaches for encouraging faculty and students to excel. *New Directions for Teaching and Learning.* San Francisco, CA: Jossey-Bass publishing.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*, pp. 223-232.

Saxe, K., & Braddy, L. (2015). *A common vision for undergraduate mathematical sciences programs in 2025.* MAA: Mathematical Association of America and Common Vision: for Undergraduate Mathematical Sciences Programs in 2025. Retrieved from https://www.maa.org/sites/default/files/pdf/CommonVisionFinal.pdf

Society for Industrial and Applied Mathematics. (2012). *Modeling across the Curriculum.* Philadelphia, PA: Society for Industrial and Applied Mathematics.

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

Stains, M., Harshman, J, Barker, M. K., Chasteen, S. V., Cole, R., Dechenne-Peters, S. E., Eagan Jr., M. K., Esson, J. M., Knight, J. K., Laski, F. A., Levis-Fitzgerald, M., Lee, C. J., Lo, S. M., Mcdonnell, L. M., Mckay, T. A., Michelotti, N., Musgrove, A., Palmer, M. S., Plank, K. M., Rodela, T. M., Sanders, E. R., Schimpf, N. G., Schulte, P. M., Smith, M. K., Stetzer, M., Van Valkenburgh, B., Vinson, E., Weir, L. K., Wendel, P. J., Wheeler, L. B., & Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science*, 30 Mar 2018, pp. 1468-1470. Retrieved from https://science.sciencemag.org/content/sci/359/6383/1468.full.pdf

Stake, R. E. (1995). *The art of case study research.* Thousand Oaks, CA: Sage Publications.

Stake R. E. (2003). Chapter 5: Case studies in Norman K. Denzin and Yvonna S. Lincoln's *Strategies of Qualitative Inquiry* (second edition). pp. 134-164. Thousand Oaks, CA: Sage Publications

Suskie, L. (2009). *Assessing student learning* (2nd edition). San Francisco, CA: Jossey-Bass.

Tagg, J. (2012). Why does the faculty resist change? *Change: The Magazine of Higher Learning, 44*(1), pp. 6-15. Retrieved from Changemag.org

Toppings, K. J. (2005). Trends in peer learning. *Educational Psychology, 25*(6), December 2005, pp. 631-645. Routledge.

Watkins, J., & Mazur, E. (2013). Retaining students in science, technology, engineering and mathematics (STEM) majors. *Journal of College Science Teaching, 42*(5), 36–41.