The Report committee for Sara Estelle Cohen Certifies that

this is the approved version of the following report:

The Calculus Concept Inventory and Teaching Methodology Reform

APPROVED BY

SUPERVISING COMMITTEE:

Supervisor: ____

Lorenzo Sadun

Co-Supervisor: ____________________________________

Mark Daniels

The Calculus Concept Inventory and Teaching Methodology Reform

by

Sara Estelle Cohen, B.S.

Report

Presented to the Faculty of the Graduate School

of the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

May 2014

The Calculus Concept Inventory and

Teaching Methodology Reform

by

Sara Estelle Cohen, M.A.

The University of Texas at Austin, 2014

SUPERVISOR: Lorenzo Sadun

CO-SUPERVISOR: Mark Daniels

Unlike fields in the humanities and social sciences, mathematics is traditionally taught through lectures in which students are expected to passively learn material. Research has shown that this didactic method leaves students with little conceptual understanding and discourages them mathematically. The Calculus Concept Inventory (CCI) is an exam which was developed to determine the impact of different teaching methodologies on students' conceptual understanding. The results have demonstrated that teaching methods which fall under the category of Interactive Engagement have the largest positive impact on conceptual knowledge. These methods actively engage students through social interactions with their peers and instructors in addition to providing immediate feedback and time for second attempts. The purpose of this report is to describe the current status of calculus reform and the ways in which the CCI is affecting mathematics pedagogy. For example, the University of Texas at Austin Mathematics Department has implemented flipping, an interactive engagement method delivering instruction on-line outside of class and bringing homework into the classroom.

Table of Contents

Chapter One: Introduction

The way in which calculus is routinely taught makes it especially apt for change. Calculus has a tradition of being very didactic in its dissemination of content through lectures as opposed to fields in the humanities and social sciences which traditionally place more emphasis on exploring concepts. In the fall term of 2010, the Mathematical Association of America (MAA) conducted a national survey of Calculus I instruction across a stratified random sample of two and four-year colleges and universities under the title Characteristics of Successful Programs of College Calculus (CSPCC). [4] The survey disturbingly showed that research universities, which teach Calculus I to more students than any other category of post-secondary institution, also seem to be doing the worst job in maintaining students' confidence in their mathematical abilities, enjoyment of mathematics, and interest in continuing with the mathematics that is needed to pursue their intended careers. [4] Calculus I is the primary gateway for most students heading into the technical and scientific fields that will drive the economy of the 21st century. The CSPCC study has revealed that the students who enter this gateway are highly motivated and consider themselves well prepared; however these students are being mathematically discouraged. It is clear that there is great need to reform our mathematics education.

The Calculus Concept Inventory (CCI) is an exam that was developed in 2005 with the intention of aiding in the rehabilitation of calculus. The CCI is used to compare the impact of different teaching methodologies on students' conceptual understanding of differential calculus. Through the analysis of these results, the goal is to identify qualities of teaching styles which lead to greater improvements in students' conceptual understanding. The idea for the Calculus Concept Inventory succeeded the achievements of the Mechanics Diagnostic Test (MDT) and the Force Concept Inventory (FCI) in physics education reform. These exams were administered before and after courses and *normalized gains* were calculated. Normalized gains measure the improvement in the class's performance as a fraction of the maximum possible improvement. The results of the FCI showed that students being taught physics using a traditional - lecture based - approach came out of the course with little increased conceptual understanding. On the other hand, Hake found that students taught with *Interactive Engagement* (IE) methods, which will be described later in this paper, were shown to develop a better grasp of general physics concepts with an average normalized gain about two standard deviations higher than the average normalized gain of traditional courses. [6]

The National Science Foundation funded the Calculus Concept Inventory development, which was done by calculus educators and a standardized test consultant. This test, similar to the MDT and FCI, has shown that students taught with interactive engagement teaching methodologies have a superior grasp of basic calculus concepts. This paper will describe the favorable qualities of multiple teaching methodologies in addition to the development and validation of the CCI; the results of which have already begun to aid in the improvement of teaching techniques and benefit mathematics education.

Chapter Two: What is a Concept Inventory?

Before going into the development of the Calculus Concept Inventory, it is important to define the term *concept inventory.* Using the definition given by the FCI in physics: concept inventories are tests of the most basic conceptual comprehension of foundations of a subject and not of computational skill. [6] These inventories do not attempt to test everything taught in a course and are quite different from exams given in typical classes. Considering that the MDT and FCI form the roots of the concept inventory industry, their methods and reasons for success will be discussed.

Over the past few decades, the FCI in physics has spurred a dramatic change in physics curricula and teaching methods. This triumph is due to a variety of factors. Firstly, physics educators have agreed that questions asked on the FCI have practical use even on problems involving computation. Faculty decided that the concepts are necessary, although not adequate, for understanding basic physics. Many studies, including Hake's ([9],[10]), showed that the FCI provides a reproducible and objective measure of how a course improves comprehension of principles, not merely how bright or prepared the students are nor what they have memorized. Hake and others who analyzed the FCI results calculated a performance measurement called the *normalized gain* $< g >$ which is given by the following formula: $< g > = \frac{\mu_f - \mu_0}{100 - \mu 0}$ where μ_0 is the mean score at the start of the class and μ_f is the mean score at the end of the class. ([9],[10]) This measures the gain in the class's performance as a fraction of the maximum possible gain.

Hake surveyed sixty-two courses and shockingly found that the normalized gain $< g >$ is essentially independent of μ_0, with a correlation of only ±0.02, and is also largely independent of instructor and text. [6] On the other hand, there is a strong dependence between the normalized gain $< g >$ and teaching methodology. Hake found that classes using traditional approaches had an average normalized gain of 0.23 (standard deviation of 0.04) and courses using interactive engagement had an average normalized gain of 0.48 (standard deviation of 0.14). [6] The substantial correlation between the normalized gain and teaching methodology is remarkable and consistent among studies. Data and analysis of these results have provided objective evidence, which convinced many educators to alter the way they teach and seek validation from the FCI. This movement in physics has been so impressive that there are now concept inventories in many other subjects including math (CCI), biology, chemistry, and astronomy. The results of the Calculus Concept Inventory have similarly demonstrated that teaching methodology is strongly correlated with average normalized gain suggesting that calculus would benefit from a comparable transformation.

A concern from educators is that the interactive engagement methodology does not sufficiently prepare students to be able to solve computational problems. There is much physics research ([9],[10]), however, that shows that students being taught with IE methods are no worse at solving such problems compared to students taught traditionally. Mazur, a Physics Professor at Harvard, developed *Peer Instruction*, which is a teaching methodology that involves much less lecturing and working examples and much more student involvement. Mazur expected a decline in success solving the computational problems at the end of the text; however, he found no difference between students solving abilities when taught using peer instruction versus with a traditional approach. On the other hand, Mazur found that students in peer instruction courses had better conceptual understanding of the material. The goal of the Calculus Concept Inventory is to see if this result also holds true in calculus in that there is a conceptual improvement and no decline in the ability to solve computational problems while using interactive engagement teaching methodologies.

Chapter Three: Interactive Engagement

More schools are becoming involved and administering the CCI to their calculus students. For this reason, there is a wealth of data on traditional courses; however, data from interactive engagement classes is more difficult to obtain. This is partially due to ensuring that the classes satisfy an independent definition of IE, which is necessary for a valid comparison. This paper will use the following definition, which was given by Hake: *Interactive Engagement* (IE) methods are those designed at least in part to promote conceptual understanding through interactive engagement of students in heads-on (always) and hands-on (usually) activities which yield immediate feedback through discussion with peers and/or instructors. [10]

Epstein found that the "immediate feedback" part of the definition was essential in that courses which favored activities as opposed to lecture did not necessarily have a high average normalized gain. [6] It is important for students to be given immediate feedback from the instructor or other students in order for them to see if their answers make sense and are consistent with the other concepts they have learned. Students need to be given time to assess their work and try again if needed. Many other studies, including those done by Shavelson [11] and Black and Williams [8], have also emphasized the value of immediate formative assessment.

Discovery Learning is another term that is often used to describe teaching methods in which students are actively engaged. Dewey, Vygotsky (1930), and Piaget

(1950) set the groundwork for this approach to learning and hypothesized that students learn best by actively constructing their knowledge through social interactions, rather than simply absorbing ideas directly. [7] Five characteristics are identified by The National Academy of Sciences as vital to discovery learning models in mathematics and the sciences. It is critical that the student: (a) be engaged by scientifically oriented questions; (b) give priority to evidence; (c) formulate explanations from evidence; (d) evaluate explanations in light of alternative explanations; and (e) communicate and justify proposed explanations. [2] Some teaching methods that fall under this umbrella are inquiry-based, problem-based, Socratic, and Moore-Method instruction. Research on discovery learning asserts the following advantages: promotes creative thinking and the development of higher level thinking skills, sees failure as a natural and often essential step on the pathway to success, engages and motivates students, enhances confidence, and develops employable skills such as problem solving, communication, collaboration, and presentation skills. [7]

Another teaching methodology that often satisfies the definition of interactive engagement is called *Flipping*. Flipping describes the inversion of expectations in the traditional lecture. Students gather the course material largely outside of class, by reading, watching videos or recorded lectures, or listening to podcasts, and when they are in class, students do what is typically thought to be homework, solving problems with their professors or classmates, and applying what they learn to new problems. [3] Students continue this process on their own outside class. The immediacy of feedback in this way enables students' misconceptions to be corrected well before they show up on an exam. According to a growing body of research as described above, this results in more effective learning.

Chapter Four: Development, Validation, and Results of the CCI

NSF funding was obtained by Epstein and Yang in 2004 for the development and validation of the Calculus Concept Inventory. Howard Everson, a standardized test consultant, joined the project to perform validation studies on the CCI and the collected student data. A group of highly experienced mathematics faculty extensively discussed and agreed upon a set of basic differential calculus concepts that the test should cover. In the spring semester of 2005, the first CCI was assembled, produced, and given to about 250 students from six different institutions as a pre-test. Despite the fact that most of these students had previously taken calculus, the test results seemed to be equivalent to random guesses. Students were given this test again at the end of the semester and there was no gain among any of the institutions. It was quite shocking that there was no gain and was thus time for a reassessment.

The panel came to the conclusion that, even though they believed the questions on the exam were basic, they were too hard for the purposes of the concept inventory. This group decided that the questions needed to be at a level which faculty would believe the answers to be trivial. In the fall of 2006, a revised version was ready and administered to about 1,100 students at fifteen different institutions in the U.S. and one in Finland. Pre-test scores were above the random guess level and there was some gain at every institution at the end of the semester. Although legitimate interactive engagement sections of calculus are hard to come by, Epstein was able to get data from Uri Treisman at the

University of Texas, from a strongly IE-based instructor at Oregon State University, and from two sections at St. Mary's College of Maryland, of clearly IE methodology. All of these IE sections showed a normalized gain between 0.30 and 0.37, which was well above the gain for the traditional courses in Epstein's study. [6] It must be remarked, however, that although these results were enticing, there was not adequate data from IE sections at that point (2006) in order to make broad conclusions. The next step was evaluating the validity of the exam, which was done by Everson and his graduate students.

Scores on exams can shed much light on students understanding if the questions are actually measuring what they are intended to. It is possible, however, that students get right answers for wrong reasons or get questions wrong due to faulty wording. Cognitive Laboratories, also called "analytic interviews", are great for detecting this anomaly and were thus helpful tools to use in the validation of the CCI. [6] Cognitive Labs are highly structured interviews in which students are asked to think out loud as they work out a problem. A very carefully designed protocol is then used to question the students in order to gain insight into their mental process. Everson and his graduate students designed such protocol for the CCI and during spring 2006 two out of twenty four questions on the exam were found to be potentially problematic and were removed.

The lack of sufficient IE data was remedied in the fall of 2008 when the CCI was administered to all fifty-one IE sections of Calculus I at the University of Michigan totaling about 1,342 students. [6] At U-M, calculus instructors, which are mostly grad students and post-docs, are provided with extensive resources for teaching their individual sections. Additionally, homework assignments and exams are created by course coordinators and are identical among all sections. New instructors, which accounted for eighteen of the fifty-one, are required to attend a pre-semester training workshop and weekly meetings to support their teaching. Experienced instructors often attend these meetings as well. In order to insure legitimacy of the CCI, it was administered on-line in a proctored lab to which no instructor had access. Although there were some incentives given, a surprising 96% of students took both the pre and post-test which were administered the first and last week of the course, respectively. The dropout rates for Calculus I at U-M were much lower than typical at other universities, which alone is an important independent result. Extraordinarily, the average gain over all fifty-one sections was 0.35, ranging from 0.21 to 0.44. Epstein and his colleagues found that the lowest gain of 0.21 at U-M was about the highest gain found among traditional calculus courses at all other schools in their study. [6] The section at U-M with the lowest gain contained twelve students who were at risk of failure. It is daunting that this section had the same gain as the section with the highest gain among the traditional sections.

All institutions in this study were given the same CCI and compared via the same analysis, i.e. the normalized gain. Computerized instruction programs have shown the same normalized gain as traditional classes, which displays the consistency of the CCI in measuring the outcome of a course. This is essential in order to determine what features actually make a difference on students' understanding of concepts. On average, the IE sections had significantly larger gains compared with the traditional sections, which is similar to the FCI in physics. Considering the fact that more than a third of the instructors at U-M were new, it appears that IE teaching methods can be learned and successful when instructors are provided with sufficient materials and resources. The FCI has shaped physics education in many ways and there is hope that the CCI will have a similarly significant impact on calculus. Universities are encouraged to use the CCI to contribute to the ongoing study and to aid in positively impacting mathematics pedagogy.

Chapter Five: Calculus Reform at UT Austin

The University of Texas at Austin Mathematics Department began transforming their mathematics education in the spring semester of 2012. Elizabeth Stepp and Jane Arledge spearheaded this shift in teaching methodology and flipped their sections of calculus, which contained about 120 students each. Stepp and Arledge assigned readings from the text and gave on-line pre-class assignments. Despite the additional work, Stepp and Arledge enjoyed flipping more than lecturing and additionally felt that students achieved a deeper understanding of the material. This spurred the decision to flip more sections of calculus and over the summer of 2012, Elizabeth Stepp began creating a series of videos to be used as part of the pre-class assignments, or learning modules, for these calculus courses. Lorenzo Sadun aided Elizabeth Stepp in the completion of the videos in Fall 2012. These videos were better integrated with the on-line homework and faculty were generally pleased with the results. Throughout 2013 more materials including videos and pre-class and post-class assignments were created and additional calculus courses were flipped. Faculty felt these courses were successful, citing higher attendance, significantly lower failure and dropout rates, and an overall perception that students gained a superior understanding of the material.

One issue with flipping calculus is finding software that is able to harmonize pre-class assignments, homework, and exams. UT is currently using a software program called Quest which provides tools to incorporate on-line multimedia content and assessments into courses. UT uses Quest to create homework assignments and exams in addition to aiding in the creation of the pre-class assignments. Although Quest has a few gaps in their material, it is currently suiting the needs of UT which allows UT to focus on improving their learning modules and creating other materials. The ideal situation would be to have shorter, more engaging videos and pre-class assignments that would interact with the user and provide review or additional problems when mistakes are made. In this way, not only would the students acquire immediate feedback during class, but they would also be engaged by their coursework and their misconceptions or miscalculations could be addressed sooner.

Additionally, UT made the conscious decision to allow professors freedom in how they teach these calculus courses, unlike U-M, which provides a teaching schedule, resources, and mandatory training. Although this allows professors to have control over their specific teaching style with the option of using provided worksheets, it means that professors need to dedicate more time in order to be properly prepared for their courses. The first sections that were flipped at UT were taught by professors who were open to educational experimentation. As UT has increased the number of courses flipped, the importance of having an instructor with a positive attitude towards trying new methods and being willing to learn them has grown. It can be difficult to persuade professors who feel they give successful lectures to try a new teaching technique. Hopefully, as UT continues to improve their methods, a new teaching culture will be built and will catch on over time.

UT is one of the first universities to implement interactive engagement methods in such large calculus courses, and has approached the project with great optimism and expectation of success. In Fall 2013, UT ran an experiment in which they flipped half of the calculus courses and used a traditional lecture style for the other half. These courses were given the same homework and tests and were additionally administered the CCI. Results will be released soon which will provide significant data, aiding in the continual study and improvement of mathematics pedagogy.

Chapter Six: Conclusions and the Future

Research consistently shows that students cannot passively receive and understand material in class. In order to truly learn and comprehend new information students must be actively engaged, whether through group work, student presentations, or other strategies. Immediate feedback is also crucial in ensuring that their ideas make sense and are compatible with previously learned concepts. Additionally, students need to be given the opportunity to try again if they do not produce the correct results. This way they are actively participating in connecting concepts and putting them to use through trial and error.

The normalized gains for interactive engagement Calculus I courses were significantly higher than those for traditional courses. The data sets were essentially two gaussians with no overlap. [6] The results of Epstein's study are noteworthy; nevertheless, we must proceed with caution. There is always the possibility that some unaccounted variable is skewing the results; however, many variables have been taken into consideration. These include class size, textbook, instructor, class time, and students' previous knowledge. More and more institutions are administering the CCI so within the coming years the additional data should provide for more validation and analysis of the results and possible effects of any lurking variables should become clearer. The results of the CCI are already starting to affect mathematics education and will continue to do so. Based on the research discussed, the author believes that mathematics programs should adopt interactive engagement teaching methodologies with the purpose of improving students' conceptual knowledge in addition to maintaining and broadening interest and engagement in mathematics.

References

1. Arnold Arons, Toward a wider public understanding of science, Amer. J. Physics 41(6) (1973), 769–782.
2. Committee on Development of an Addendum to the National Science Education Foundation Standards on Scientific Inquiry, (2000), Inquiry in the national science education standards, In Inquiry and the National Science Education Standards: a guide for teaching and learning, pp. 13-38, Washington, D.C.: National Academy Press.
3. Dan Berrett, How ‘Flipping’ the Classroom Can Improve the Traditional Lecture, The Chronicle of Higher Education, 2012.
4. David M. Bressoud, Marilyn P. Carlson, Vilma Mesa and Chris Rasmussen, The calculus student: insights from the Mathematical Association of America national study, International Journal of Mathematical Education in Science and Technology (2013), 44(5), pp. 685-698, DOI:10.1080/0020739X.2013.798874.
5. I. Halloun, D. Hestenes, R. Hake, and E. P. Mosca, Force Concept Inventory, 1995, <http://bit.ly/b1488v>.
6. Jerome Epstein, The Calculus Concept Inventory – Measurement of the Effect of Teaching Methodology in Mathematics, AMS, 60(8) (2013), pp. 1018-1026.
7. Mark Daniels, Discovery Learning, Accepted for Sage Publication: Encyclopedia of Educational Theory and Philosophy, 2013.

8. Paul Black and Dylan Wiliam, Inside the Black Box: Raising Standards Through Classroom Assessment, Phi Delta Kappan, 80(2) (1998), pp. 139-148, <http://www.setda.org/toolkit/nlitoolkit2006/data/Data_InsideBlackBox.pdf>.
9. Richard Hake, Interactive engagement methods in introductory mechanics courses, 1998, <http://bit.ly/aH2JQN>.
10. Richard Hake, Interactive engagement versus traditional methods: A six-thousand student survey of mechanics test data for physics courses, American Journal of Physics, 66 (1998), pp. 64-74, <http://bit.ly/9484DG>.
11. R. J. Shavelson, Formative assessment, Special issue of Applied Measurement in Education, 2008, <http://stanford.edu/dept/SUSE/SEAL/>.