Summer 8-22-2020

Use of Statistics Boot Camps to Encourage Success of Students with Diverse Background Knowledge

Justin C. Willis
University of Maine, justin.willis@maine.edu

USE OF STATISTICS BOOT CAMPS TO ENCOURAGE SUCCESS OF STUDENTS WITH DIVERSE BACKGROUND KNOWLEDGE

By

Justin C. Willis

B.S University of Maine, 2018

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Teaching

(Generalist Concentration)

The Graduate School

The University of Maine

August 2020

Advisory Committee:

Franziska Peterson, Ph.D., Assistant Professor of Mathematics Education, Advisor

Brett Ellis, Ph.D., P.E., Associate Professor of Mechanical Engineering Technology

Timothy Boester, Ph.D., Assistant Professor of Mathematics

USE OF STATISTICS BOOT CAMPS TO ENCOURAGE SUCCESS OF STUDENTS WITH DIVERSE BACKGROUND KNOWLEDGE

By Justin C. Willis

Thesis Advisor: Franziska Peterson

An Abstract of the Thesis Presented

in Partial Fulfillment of the Requirements for the

Master of Science in Teaching

(Generalist Concentration)

August 2020

This study seeks to analyze the effects of targeted interventions on students with varying statistics background knowledge. These interventions include remediating assessment homework assignments, and Maine Learning Assistant (MLA) led statistics boot camps. These interventions were completed in an undergraduate Lean Six Sigma course, where students initially had a wide variety of prior statistics experience. This large dispersion of background knowledge levels is paralleled in many STEM entry-level courses. Data collected about student participation in these interventions and their later success on exams in the course were analyzed using General Linear Model protocol to determine if any intervention created a statistically significant change in student success measures. Several models were run, each concentrating on a particular statistics background knowledge concept addressed in the boot camps and essential for course success. Examination success rates were found to have increased significantly from the cohort without

these interventions (2016) to the cohort with these interventions (2018). This improvement was maintained with the second cohort with the interventions in 2019. The statistics backgrounds of the 2016, 2018 and 2019 cohorts were not found to be significantly different from each other after analyzing their reported backgrounds and major demographic. However, no strong singular effect was found on student success through General Linear Model Analysis. Further data collection of student participation and success measures is encouraged in subsequent course offerings, to enhance the chance of detecting subtler intervention effects on student success. Qualitative data from student and MLA interviews may also be beneficial to see how perceptions of statistics and teaching are influenced by the interventions.

ACKNOWLEDGEMENTS

I would like to thank my thesis committee for insight and time spent refining this document into what it is, and the University of Maine Department of Mechanical Engineering Technology for their support during this process. I would also like to thank my family and teachers for encouraging me to go get a degree in "something useful". When I figure out what that means, I will be sure to tell them.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Statistical literacy – the ability to analyze sampled data, draw inferences about the underlying population, and assess the veracity of inferences by others – is a valuable mathematical skill employed in most people's everyday life. Whether realized or not, people in today's society are dealing with data all the time in places like news broadcasts, business reports, and advertisements. Data is even more prevalent in the science, technology, engineering, and mathematics (STEM) disciplines and occupations, where statistical analysis results may be the primary evidence to support a conclusion. Although 72% of STEM workers in 2018 reported seeing more data than ever before, only 45% of STEM workers felt prepared to analyze and draw inferences from data (Qlik, 2018).

Despite the importance of statistical literacy, improving statistical literacy faces two major impediments at the high school level. First, statistics educators are poorly prepared to teach statistics. For example, a majority of mathematics teachers listed statistics as the subject they were least comfortable teaching (Lovett & Lee, 2017). Second, statistics literacy is rarely emphasized, which manifests as a lack of statistics skills demonstration prior to graduating high school. For instance, although Maine requires 3 years of mathematics for high school graduation, almost every district satisfies the 3-years requirement with two years of algebra and one year of geometry. This lack of statistical education is further exacerbated as measuring statistics literacy on conventional standardized tests is difficult (Iddo Gal & Garfield, 1997), meaning that most first-year college students have a minimal understanding of statistical literacy skills. It is entirely possible students may not have had formal statistics teaching at all within the secondary level.

The lack of statistics education in high school means that collegiate courses serve as introductory statistics courses for most collegiate students. Collegiate courses often focus on parameterization of central tendency (*e.g.*, mean, median, mode), parameterization of dispersion (*e.g.*, range, standard deviation), and drawing statistical inferences about the underlying population, (*e.g.*, calculating a confidence interval for population mean). Other inferences can include determining how likely it is that a data point is part of a population and determining how likely it is that two data sets share the same average. These inferences can be useful when determining if a change in a variable created a meaningful change in the resulting output.

While these introductory collegiate statistics courses do exist, they are not required for all student majors, even if the students in that major could possibly benefit from the introductory statistics knowledge. Thus, college-level student cohorts will often have a large dispersion in prerequisite statistical literacy, even within a given major or field. The dispersion in a cohort's statistical literacy presents a question for educators trying to incorporate statistics concepts in their courses. Educators need to determine pedagogical approaches for teaching a collegiate-level statistics course to a cohort having a large dispersion in prerequisite statistical knowledge. These pedagogical approaches may also apply by extension to student cohorts with a wide dispersion of background knowledge in STEM fields.

A conventional approach is to assess student prerequisite knowledge and partition the cohort into two (or more) courses – an introductory course and a non-introductory course (Weiss & Belland, 2016). The introductory-course cohort would then advance to the non-introductory course in a later semester. While appropriate and useful for large-enrollment and frequently offered course sequences such as Calculus, this "assessment and partitioning" approach presents at least three disadvantages for infrequently offered, small-enrollment courses. First, this approach may decrease enrollments, particularly for upper-level electives taken near the end a student's education. Second, it disproportionately impacts students underrepresented in STEM disciplines that may have had fewer opportunities to develop statistics knowledge in precollege education. Third, this approach also causes logistical problems by being too much of a faculty credit hour strain to be viable, given the smaller and less frequent nature of this course.

A second approach could be to simplify the course such that the course includes all of the introductory content required for the least-prepared student to succeed in the course. This “simplify” approach is flawed in that students entering with adequate pre-requisite knowledge are penalized by having to review already-mastered concepts. The disadvantages of both the assessment and partition approach and the simplify approach are sufficient to warrant an alternative.

The proposed approach, which maintains the rigor necessary for the course, is to deliver the course at the standard pace and instructional goals but provide tailored instruction on required background knowledge for those that need it. While this will require additional resources, the additional time and effort required will not be equal to delivering an entirely separate course. This allows students to all be enrolled in the same course, and arrive at the same student outcomes, but come to the course starting with different levels of background knowledge. While the previous two approaches may not result in significant learning gains themselves, the advantages of this “tailored instruction” approach do not automatically guarantee its value. This study endeavors to detail such an approach and investigate the demonstrated student outcomes, calibrating for dispersion in statistical background knowledge.

Seeing this split of statistical knowledge within the class MET 440, a course in using “Lean Six Sigma” principles to improve engineering and manufacturing processes, the instructor was presented with the task of measuring the gap in knowledge, determining which students needed more instruction in prerequisite statistical literacy skills, and addressing students’ gaps with directed instruction and practice. Ideally, this would be addressed by covering all the prerequisite content and providing practice using all of the skills necessary for success in the course with the whole class, but content and certification restrictions prevented this from being viable. This led the course instructor to develop a plan involving a screening examination for all students, and extracurricular instructional sessions and practice opportunities for students who were found to have statistics literacy skill difficulties. These sessions framed as “statistics boot camps” set out to provide the required skill instruction and practice in a fashion that allowed students from all statistical backgrounds an opportunity for course completion and success. These sessions were delivered and presided over by Maine Learning Assistants, senior undergraduates who had succeeded in the course in a previous iteration and were enrolled in a teaching methods seminar concurrently.

The ability to address differences in student background knowledge within a course could have broad impacts. First, many introductory collegiate-level STEM courses have cohorts that include large dispersions in prerequisite knowledge, therefore this study may help inform how to address this issue in other STEM courses. Second, dispersions in cohort prerequisite knowledge exist in non-introductory STEM courses due to overtaxing cognitive capabilities of lesser-prepared students in the introductory STEM courses. Addressing this early on in student trajectories may help alleviate attrition of STEM student enrollment, as it prevents the gaps in student background knowledge from compounding to the point that the student has extreme difficulty succeeding in courses. Third, by addressing cohort dispersions in introductory and non-introductory STEM courses, conclusions drawn from this research may help address the persistent underrepresentation of females and minorities in STEM fields.

The remainder of this thesis is organized as follows. Chapter 2 reviews the literature regarding the prominence of statistical literacy in students entering STEM disciplines, previous strategies used to provide support for students with varying background knowledge levels in introductory STEM courses, and specific challenges the course in question presents. Chapter 3 provides details about the three course offerings studied, the data collection, models and analytical methods to interpret the data, and research questions the data analysis seeks to answer. The results of model analysis are presented in Chapter 4, with detail for each model analyzed during the study. Chapter 5 synthesizes results from data analysis, introduces conclusions that can be reached by analysis of the data, and includes recommendations for further study. Selected examples of questions used in data collection at different stages of data collection are provided in the Appendix.

CHAPTER 2: LITERATURE REVIEW

Theory Basis for Study

Theoretical foundations for this study include the application of social cognitive theory of learning, towards developing student ability to demonstrate statistical reasoning through demonstration and guided practice. The method of guided practice aims to assist students develop a clearer cognition of the setups required in the process and helps to address knowledge "gaps." Students were screened with a Concept Inventory of prerequisite statistics concepts, which allowed students to develop a clearer idea of their self-efficacy, helping to inform their learning and practice decisions. By participating in lessons concurrent to the general classroom instruction, students saw models, demonstrated skills, and received immediate instructor/model feedback.

Providing students with the required skills when needed strengthens the cognitive connections developed between the demonstrated statistics analysis skills and the engineering outcomes desired by the students (Acee & Weinstein, 2010; Schunk, 2012). This strong behavior-value connection leads students to develop increased motivation to develop the skills needed for the course, due to the increased proximity between the goals set and the values of the students (Schunk, 2012). The methods used for generating these stronger cognitive connections and skills and motivations are also changed and directed based on student feedback, through their observable behavior and interaction with the class environment. This allows a student's cognition and behaviors to influence external models, in accordance with Bandura's model of triadic reciprocity in social cognitive theory (Bandura, 1986). Bandura's model of reciprocal determinism also expresses that the class environment and skills demonstrated by the students can influence and be influenced by a student's observable behavior (Bandura, 1986).

Statistical Literacy

Statistical literacy is defined as the ability to responsibly work with and analyze data, and to generate responsible conclusions from the data provided (Wallman, 2006). With these conclusions, students then can hypothesize how to apply these conclusions to new and similar data and make recommendations towards real world implications or changes (I Gal, 2004). This process requires basic mathematical knowledge of analysis properties, a specialized statistical knowledge to make sense of the data provided, and data context (Sharma, 2017). This process is highly desired by employers and workplaces, especially as more and more businesses are using data in increased capacities to make decisions. Seventy-two percent (72%) of worldwide working professionals surveyed say they are working with more data in making decisions then they did three years ago, while 55% feel they have inadequate education and tools to make decisions based on this data (Qlik, 2018).

Statistics and the common core state standards for mathematics

Working with statistical processes and data has a strong basis in the Common Core State Standards for Mathematics. Starting in the 6th grade standards, students are expected to think about questions in a statistical manner; analyze central tendency, range, and extremes; and draw inferences. Sixth-grade students also visualize data with bar graphs, line graphs, and histograms, and choose appropriate visualization methods based on what you are trying to convey (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). This understanding continues in the higher grades of middle school towards a qualitative understanding of how a sample is related to its representative population. Within the "Interpreting Categorical and Quantitative Data" heading of the Statistics and Probability section of the high school standards, expectations for students regarding using measures of central tendency (mean and median) and dispersion (standard deviation, data quartiles and range) to analyze data are expressed (National Gov., 2010). Expectations for students to quantitatively generate a range for a population mean based from a selected sample's mean, sample size and standard deviation are also set for high school students in the "Making Inferences and Justifying Conclusions" heading of the same section (National Gov., 2010).

Challenges for implementing statistics standards in 6th – 12th grade

In practice these standards are rarely met. A study of US students where middle and high school students were piloting a new measure of statistical literacy, The Levels of Conceptual Understanding in Statistics (LOCUS), found that students at both the middle school and high school level showed a marked decrease in performance on data analysis questions, as opposed to data collection and question formulation questions (Whitaker et al., 2015).

Further showing the universality of these challenges, Yolcu (2014) found that advancing in grade level was an unreliable predictor of data literacy in a study of Turkish middle-school students between 6th and 8th grades. Yolcu (2014) hypothesized this was due to the "spiral" nature (*i.e.*, subjects were repeated in each grade level in a cyclic manner, with subject manner depth increasing with each grade) of the Turkish state curriculum and that these students were preparing for an 8th grade state exam that was highly arithmetical and algebraic in nature

While students in Turkey are admittedly not subject to the Common Core Standards for Mathematics, and such a spiral curriculum is not mandated for use in the US, parallels still exist. US students are often prepared strongly for highly algebraic and function-based mathematics assessments in 8th grade and high school in their mathematics courses (Daun-Barnett & St. John, 2015). Statistical questions are rarely included on these multiple choice assessments, and "statistically framed" questions that are present are often proceduralized to the point where they measure mathematical reasoning as opposed to statistical understanding (Iddo Gal & Garfield, 1997). A survey of preservice mathematics educators also found that the majority of teachers were most confident in their ability to teach algebra, while 63% reported they were least confident in teaching statistics (Lovett & Lee, 2017). Teacher perceived self-efficacy has been shown to alter the presentation and breadth of concepts covered in a mathematics course, much like how student self-efficacy affects student performance on these subjects (Schunk, 2012). Teachers are likely to spend less class time on subjects with less test representation and state mandated requirements, such as statistics. These subjects often tend to be placed at lower priority levels as well (Daun-Barnett & St. John, 2015). This may be a reason why University of Maine (UMaine) undergraduates demonstrated only marginal improvements in graphical analysis and extrapolation compared to Maine 8th graders (Bragdon, 2014). Lacking statistical literacy, students are neither able to draw conclusions from data nor design and evaluate experiments of their own (Rumsey, 2002).

In conclusion, the existing standards for developing statistical and graphical literacy in middle- and high-school grade levels have not yet adequately influenced real-world K-12 instructional practice. Therefore, introductory statistics skills are frequently taught in college-level courses.

Pedagogical challenges in introductory statistics courses

Success in statistics courses depends upon different skills than other mathematics courses. For example, Johnson and Kuennen (2006) concluded that science ACT subscores more reliably predicted student performance in an introductory statistics class than mathematics ACT subscores. The science portion of the ACT was hypothesized to allow students to demonstrate skills of reading graphs and charts, and evaluating hypotheses. These skills were found more beneficial to students in this statistics course than skills with purely mathematical operations, like calculus (Johnson & Kuennen, 2006). Development of these statistics-specific skills has been encouraged in introductory statistics courses by use of student-centered active learning instruction methods (Bateiha, Marchionda, & Autin, 2020), and through provided memory mnemonics provided to students (Mocko, Lesser, Wagler, & Francis, 2017). These methods are also commonly used to encourage student engagement with course material in other courses, particularly introductory STEM courses such as basic biology and precalculus. Effectiveness of voluntary student tutoring sessions led by past students has also been shown in increasing student retention and course success (Batz, Olsen, Dumont, Dastoor, & Smith, 2015).

Challenges for undergraduate engineering courses

Undergraduate engineering courses, in particular the introductory “barrier courses” in introductory physics and calculus, have been shown to have a disproportionate effect on the persistence of entering STEM students in their majors, and in their persistence and completion in college in general (Baker-Ward, Mohr, Dietz, Forrest, & Felder, 1993). Performance in a first college mathematics course, either at the calculus or precalculus was a significant predictor of student retention and success in engineering, and that the success of the student could be predicted to some extent from just their first semester of introductory courses (Budny, Bjedov, & LeBold, 2002). This study by Budny et al. (2002) also showed that 84% of students who left engineering programs did so before their program became discipline specific, which often happens in the sophomore year. This shows that difficulty in introductory mathematics courses and knowledge gaps in fundamentals of mathematics reasoning can prevent students from succeeding in an upper level engineering topic course such as MET 440, and may even prevent students from reaching MET 440 in the first place.

MET 440: Lean Six Sigma

Six Sigma (sometimes stylized as 6σ) is a method of process measurement, analysis, and improvement commonly used in engineering, manufacturing, medical services and business. The name Six Sigma derives from the goal to have the mean response be at least six standard deviations from the nearest specification limit – lower or upper. Here, the lower and upper specification limits are defined as the minimum and maximum values, respectively, required by the customer. If the six standard deviation goal is met, the process will produce a defect approximately 3.4 times per million opportunities (Ali & Ahmed, 2016).

The capability for a process to meet this goal is parameterized by capability process indices, *i.e.*, calculations of variation and process capability that then inform process improvement decisions. Common strategies to increase process capability include determining rational specification limits, centering the process mean between the control limits, and decreasing the variation in the process. These improvement strategies roughly ranked by difficulty are then analyzed by a Six Sigma team and recommendations are made to the company based on their capability findings. This process is visualized in the Define, Measure, Analyze, Improve, and Control (DMAIC) cycle (Quality Council of Indiana, 2014).

Statistical challenges for MET 440

The Analyze step of DMAIC and the capability indices within the Analyze step requires a thorough understanding of sample mean and sample standard deviation. This understanding is used both to develop measurements of a sample and to understand what our sample can tell us about the general population. For instance, a sample mean does not actually give us the true mean of the overall process; rather, the sample mean can be used to calculate a confidence interval for the true mean (Quality Council of Indiana, 2014). For reasons explained above, one could reasonably expect a subpopulation of the cohort that are unable to demonstrate understanding of these prerequisite concepts. Students in this subpopulation would reasonably include those who had not enrolled in a college statistics course of any type prior to their enrollment in MET 440. Without this understanding of these more basic statistical concepts, any analysis of data performed by students as part of a Six Sigma process would be purely instrumental in nature, hindering the student's ability to assess their own work for validity and make logical conclusions from their results (Skemp, 1978).

Selection of Targeted Interventions for Undergraduate Student Success

For STEM courses that require a baseline of statistical knowledge, students are at a distinct disadvantage if they have not had a course that develops the ability to test hypotheses of collected or provided data, whether those courses are taken in high school or college. This is the case with the course analyzed in this study, MET 440: Lean Six Sigma. Many introductory STEM courses in undergraduate disciplines share this difficulty, where students of diverse academic and mathematics backgrounds are presented with the same expectations in a course, despite their varying baseline knowledge (Pyzdrowski et al., 2013). The challenge of the instructor then is how to structure the course to address and remediate baseline knowledge deficiencies, while continuing to meet the final course objectives and keep students who already have the background knowledge engaged in the course. Sub-challenges include how to determine which students are going to have difficulty due to lack of baseline knowledge, how to best address this knowledge deficit in a timely manner, and how to measure the efficacy of the chosen intervention (Reisener, Dufrene, Clark, Olmi, & Tingstrom, 2016). Although a difficult proposition, effective implementation can improve probabilistic and mathematical reasoning, which in turn can improve student success in more advanced statistics courses (Primi, Donati, & Chiesi, 2016).

Screening of students for possible qualification for an intervention is usually a first step in a targeted intervention process. In some cases, it is possible to recognize the different populations in a class due to more demographic means. For instance, a student's major may be a possible indicator of whether or not a student has likely taken a college statistics course before. Therefore, some of the selection for intervention could be initiated upon course entry. However, this process can be honed and personalized by use of an entry screening examination. Such examinations were shown to have moderate predictive ability of course success in introductory engineering courses in mathematics, and moderate prediction of placement in a first mathematics course of the appropriate level (Hieb, Lyle, Ralston, & Chariker, 2015). Hieb et al. (2015) also found other factors (*e.g.*, time and environment management, internal goal orientation, and test anxiety) influenced eventual student success. Our study of MET 440 is in a unique position to examine these factors as well. Unlike the majority of courses studied which are introductory and compulsory in nature for students, MET 440 is a technical elective for students at or above their junior level of their engineering career. This allows us to examine the effect of our interventions on students who have presumably developed effective study and time management habits and have aligned their internal goals with those of the course, as evidenced by the student electing to take this course voluntarily.

A method of intervention found in the literature most commonly is the process of adaptive assessment, where assignments and online examinations are altered in real time based on a student's responses (Liu, McKelroy, Corliss, & Carrigan, 2017). By changing the assessments to be adaptive in nature and develop from the student's baseline understanding at the present towards the course objectives, the assessments hope to build up student understanding through feedback and engaging cognitive processes of learning. A study of this method of intervention among pharmacy students showed that this method had limited to no bearing on student improvement, depending on subdiscipline studied, with chemistry showing the most improvement (Liu et al., 2017). This led to the conclusion that these sort of programs have to be implemented very specifically, designed carefully and require strong engagement with students to succeed (Liu et al., 2017).

Another method used in studies at the undergraduate level is remediating assessment, where students are allowed to retry assignments and homework questions they had missed to receive credit (Howard, Meehan, & Parnell, 2019). This method of intervention ensured that students were able to complete homework assignments – a random selection from a pool of questions – as many times as they desired until they received the score they sought. In the literature this was found to help understanding of moderate to high-achieving students, those with small and dispersed understanding gaps, but was not found to help students with more substantial gaps, due to them not understanding the prerequisite knowledge at the level where they could understand the question or their errors (Howard et al., 2019). These students may need more direct and active assistance with the course material, which has been shown to be possible with learning assistants.

Learning Assistants

Learning assistants (sometimes called LAs) are undergraduate students that assist the main course instructor in facilitating active learning in their classrooms (Otero, Pollock, & Finkelstein, 2010). Learning assistants are implemented by numerous colleges and courses, often in introductory STEM courses. For instance, the University of Maine implements their learning assistant program through Maine Learning Assistants (MLAs). However, the emphasis on active learning classrooms, communication between instructor and learning assistant, and development of the learning assistant's teaching strategies are constant characteristics across programs. Learning assistant programs have been shown to increase student engagement in courses, to provide an accessible peer for students to ask questions comfortably, and provide a student perspectives on course objectives ("Learning Assistant Program: Faculty Resources," 2020)

Effect measurement of selected interventions

Measuring student response to the interventions is the next challenge present in implementing a targeted intervention program. This is necessary to get information on student progress, recommend students for further intervention if necessary and to analyze the efficacy of intervention methods when informing their later use. Some research has been conducted into the use of brief experimental analysis (BEA) of possible interventions and their efficacy in targeted mathematics interventions (Reisener et al., 2016). In this process, interventions are tried for a brief period with the resulting gains then used to extrapolate future potential gains over the full intervention timeline, in an effort to find the best intervention for the student. This process was shown to be an effective predictor of which method would produce the best results in struggling mathematics students in the 10-12 year age range (Reisener et al., 2016), though in the way applied would be more useful in an individualized intervention, rather than the group intervention being implemented in this study. However, the concept of providing feedback to students in a regular and timely fashion regarding the effects of the intervention on their performance was used, as opposed to a school or specialist making this decision for them. Like in Reisener's study, comparative new ground in this study examines how the students themselves use this information to make choices regarding intervention participation.

Having students making this decision for themselves as responsible adults highlights more than normal a required facet of interventions for students: the effect of an intervention on the students perceived value and utility of the course material and the valued consequences of completing tasks in the course (Harackiewicz & Priniski, 2018). For an intervention to be appreciated by the students, the tasks of the intervention need to be related to an outcome the students value. In this study, this is attempted by connecting statistical concepts required directly to Lean Six Sigma engineering process improvement steps. Sometimes this value alignment step alone is able to encourage student engagement and eventual course success. Value appraisal intervention in a statistics course was found to increase course task value and demonstration of interest, with some improvement in course grade as well (Acee & Weinstein, 2010). If students see increased success on course measures like grades due to the intervention, it can serve as a positive value to the student, encouraging further engagement with the intervention and creating a feedback loop.

CHAPTER 3: METHODS

In order to examine the effect of MLA-led boot camps and remediating assessment on learning outcomes of students having disparate prerequisite knowledge, learning outcomes were assessed in MET 440 "Lean Six Sigma" during three offerings: Fall 2016, Spring 2018, and Spring 2019. MET 440 was offered for the first time in Fall 2016, which serves as the control offering; the second and third offerings in Spring 2018 and Spring 2019, respectively, included the two targeted interventions.

Fall 2016

The first offering in Fall 2016 served as the control offering and included neither MLA-boot camp nor remediating assessment interventions. Day to day work consisted of assigned readings from the *Certified Six Sigma Green Belt Primer* (Quality Council of Indiana, 2014), *The Six Sigma Way Team Fieldbook: An Implementation Guide for Process Improvement Teams* (Pande, Neuman, & Cavanagh, 2007), and *The Lean Six Sigma Pocket Toolbook: A Quick Reference Guide to 100 Tools for Improving Quality and Speed* (George, Rowlands, Price, & Maxey, 2005). Reading length varied from week to week, with the minimum being 29 pages and the maximum being 158 pages. Readings for a week were mostly out of one of the three texts, with 4 of the 12 weeks having passages from both the *Certified Six Sigma Green Belt Primer* and the *Six Sigma Way Team Fieldbook*. Multiple-choice quizzes on the assigned readings would then occur the following class session. Quizzes started at the beginning of the semester as individual response with peer grading. Later in the semester, they moved to individual response with group discussion. The course met once a week for a 3-hour session with a 10-minute break near the middle of the class time. Quizzes occurred either at the beginning of class or upon returning from the mid-class 10-minute break.

The course consisted of 26 students: 21 from Mechanical Engineering Technology (MET), and 5 from Chemical and Biological Engineering (CHE/BEN). These students were predominantly junior and senior undergraduate-level students, who chose to take the course as an elective for their degree program. Students on the whole had not heard from their peers regarding the course objectives and difficulty, as this was the first time a course resembling this one had been taught at this university.

Two one-hour-duration multiple-choice paper-based exams were delivered near the middle and the end of the semester, with responses and scores anonymized before data analysis. Other data recorded were scores from the comprehensive, four-hour paper-based final exam, completed during finals week. All data collection conducted during the Fall 2016 offering was from activities conducted within the scope of the course, therefore subjects did not incur any risk that was not already incurred by deciding to enroll in the Lean Six Sigma course.

Spring 2018

The second offering of the course met twice a week for a 75-minute session each and was the first offering to institute the Lean Six Sigma Concept Inventory (LSS CI), MLA-led student boot camps, and Blackboard-based homework assignments. The course consisted of 38 students: 21 MET students, 7 CHE/BEN students, 3 Mechanical Engineering (MEE) students, 6 practicing engineers, and 1 Chemistry (CHY) student. As in Fall 2016, these students were predominantly junior and senior undergraduate-level students, who elected to take the course as an elective for their degree program. Due to the similarities between the two majors' requirements regarding statistics courses, CHB and BEN students were combined for statistics background analysis. Statistics backgrounds and majors of the 2018 cohort are below in **Table 1**. Students in the Did Not Complete (DNC) category did not complete the concept evaluation.

Table 1

Students' major and statistics background breakdown in 2018

Major	No Stats	Some HS	HS	Some Col	Col	DNC	Total
BEN/CHE	1	0	0	3	4	0	8
Grad	0	0	1	1	4	0	6
MEE	0	0	0	0	3	0	3
MET	4	4	3	1	6	3	21

The data in **Table 1** were generated by students in the first week of the course as part of the LSS CI, a screening examination created by the course instructor to determine their prerequisite statistical literacy. The LSS CI consisted of 2 background questions in which students self-reported levels of completion in statistics and mathematics, 18 multiple-choice questions assessing student prerequisite knowledge (16 questions relate to statistics, 2 questions relate to other Lean Six Sigma skills), and 18 questions for students to state their confidence in their answers. To encourage honesty in assessment of prerequisite knowledge, students were notified and awarded an LSS CI grade based upon completion, not correctness. For example, if a student submitted the LSS CI assignment by the due date, the student was awarded 100%. Alternatively, if a student failed to submit the LSS CI by the due date, the student was awarded a 0%.

Twenty-four of the 38 students scored an 8 or below and were notified they were to attend the statistics boot camps. Students earning greater than an 8 were encouraged to attend, but not required to attend. These boot camps were facilitated by the course's Maine Learning Assistant, or MLA.

Maine Learning Assistants

Maine Learning Assistants (MLAs) are undergraduate students who have completed the course they are helping to instruct with success (usually a B grade or better) and either have completed or are taking a 1-credit teaching seminar in introductory teaching strategies through the UMaine-Orono course system. MLAs assist professors in different ways in each course, though they are not authorized to assign grades to student work. In MET 440, one of the main tasks of the MLA was to administer 30 minute “boot camp sessions” that provided guided practice and modeling of introductory statistics processes and concepts employed in the course in its engineering optimization and analysis work.

This small group targeted intervention, happening in parallel but separate from the general class, is a common strategy employed in Tier 2 of the Response to Intervention (RTI) intervention model (VanDerHeyden, 2015). This model is used commonly to address student difficulty in the K-12 system. However, literature drawing conclusions on this method at the collegiate undergraduate level was not identified.

Boot Camp Structure

The boot camps consisted of seven 45-minute sessions, starting January 29th of the semester and running twice a week (except on school holidays) until February 26th. When in session, boot camps immediately followed the lectures in the same room. As the course’s official meeting time was scheduled for a 2-hour session, students did not have time conflicts. This was done to ensure all students had the ability to stay for boot camp if they so wished. Attendance was taken at each boot camp in order to determine the effect of boot camp attendance. Although students earning a score of 8 or lower on the LSS CI were notified to attend boot camps, no detrimental consequences (*e.g.*, loss of points in course) were administered if these students did not attend boot camps.

Each boot camp occurred immediately after a lecture, after those who did not wish to remain for boot camp departed the classroom. The MLA gave a 10-15-minute explanation of a certain concept. Students then exercised their newly acquired knowledge via a set of MLA-guided example problems on a worksheet for the next 10 to 15 minutes. A second worksheet containing similar applications of the same topic followed. The students completed problems on the 2nd worksheet within 2- to 4-person peer groups, with the MLA available to answer questions from the group. This group work and discussion occupied the remaining 15 to 20 minutes of the boot camp time in most cases and students could take both worksheets with them as references for later use.

The seven boot camps emphasized four concepts: (1) Central Tendency, (2) Dispersion, (3) Z-chart and t-chart comprehension, and (4) simple and compound probability. Information on student difficulties, particularly difficulties faced by students having minimal statistical literacy, found in the Fall 2016 student cohort informed the choice of the four concepts emphasized in the boot camps.

The reading assignments for the course in the Spring 2018 cohort followed the same general order through the same readings listed above for the Fall 2018 cohort, however, the previously administered written multiple-choice paper quizzes were replaced with multiple-choice assignments administered via Blackboard. Assignment questions were randomly selected and randomly sequenced from a question pool having more questions than required for a student to complete the assignment. The number of questions in the question pool was slightly greater than number of questions required for a student to complete, *e.g.,* 25 questions in the question pool with 20 questions randomly selected for the assignment. Students were permitted to submit each assignment as many times as desired with their highest score recorded as their assignment score. Accordingly, each attempt resulted in a slightly different set of problems.

During class time over the semester, two one-hour-duration preliminary exams were administered via Blackboard. Responses and scores were anonymized before analysis. The first preliminary exam, Exam 1, consisted of 25 multiple choice questions was administered in the 6th week of the course after teaching the first six boot camps. The seventh boot camp was taught on the same day as and immediately after Exam 1. Exam 2 was administered in the 12th week of the semester, well after the 7th boot camp was taught. A comprehensive, 4-hour Final Exam was completed during finals week via Blackboard in the classroom. To ensure academic honesty, all three Blackboard-based exams were proctored in the classroom. Data were collected and protected under the general liability waiver provided to and agreed to by all students enrolling in courses with Maine Learning Assistants.

Spring 2019

With identical interventions and assessments as the Spring 2018 offering, the Spring 2019 offering met twice a week for two 75-minute sessions with a day between them and the boot camp meeting beginning immediately after class for 45 minutes. Further, the Spring 2019 offering included identical data collection processes in accordance with the Maine Learning Assistant course liability waiver, which was accepted by all students.

The course cohort consisted of 28 students, including 14 from Mechanical Engineering Technology (MET), 1 graduate student, and 13 from Chemical and Biological Engineering (CHB/BEN). As with the previous two cohorts, these students were predominantly junior and senior undergraduate-level students, who elected to take the course as an elective for their degree program. This brings the total number of Mechanical Engineering Technology and Chemical and Biological Engineering students in courses with the MLA-led boot camp intervention to a total of 35 and 21, respectively. Due to the similarities between the two majors' requirements regarding statistics courses, CHB and BEN students were combined for statistics background analysis. Statistics backgrounds and majors of the 2019 cohort are below in **Error! Reference source not found.**. Students in the Did Not Complete (DNC) category did not complete the concept evaluation.

Table 2:
2019 Major and Statistics Background Breakdown

Major	No Stats	Some HS	HS	Some Col	Col	DNC	Total
BEN/ CHE	0	1	0	2	8	2	13
Grad	0	0	0	1	0	0	1
MET	5	2	2	0	2	3	14

Development of Research Questions

Research question development considered the availability of data from the three retrospective offerings, and what types of questions could be answered. Measuring student knowledge is extremely challenging, as a clear inference from test question success to subject knowledge can be very difficult to prove. Presentation of the testing can affect the outcome independent of student knowledge. For this reason, the selected inputs and outputs were measures of course outcome success and engagement. Ideally, the research questions should also help to decouple the cohort from the outcome, *i.e.*, were changes in student outcomes due to pedagogy or to the cohort?

Research Questions

- How did student outcomes of demonstrated statistical knowledge change during the three offerings?
- What differences exist in students' background statistical knowledge within the three offerings?

- What effects did statistics boot camps and remediating assessments have on student statistics knowledge outcomes in the 2018 and 2019 course offerings?

Assessing student outcomes in each course offering

Student responses in Exam 1 quantified student outcomes. Exam 1 is the most reasonable choice because Exam 1 occurred after six of the seven boot camps and contained boot-camp related topics. Apart from a paper-based format in 2016 and a Blackboard-based format in 2018 and 2019, Exam 1 questions were identical for the three course offerings. Exam 1 scores and averages were compared between the three course offerings via a one-way ANOVA using Minitab v19 (Minitab, 2019)

Assessing student statistics background in each course offering

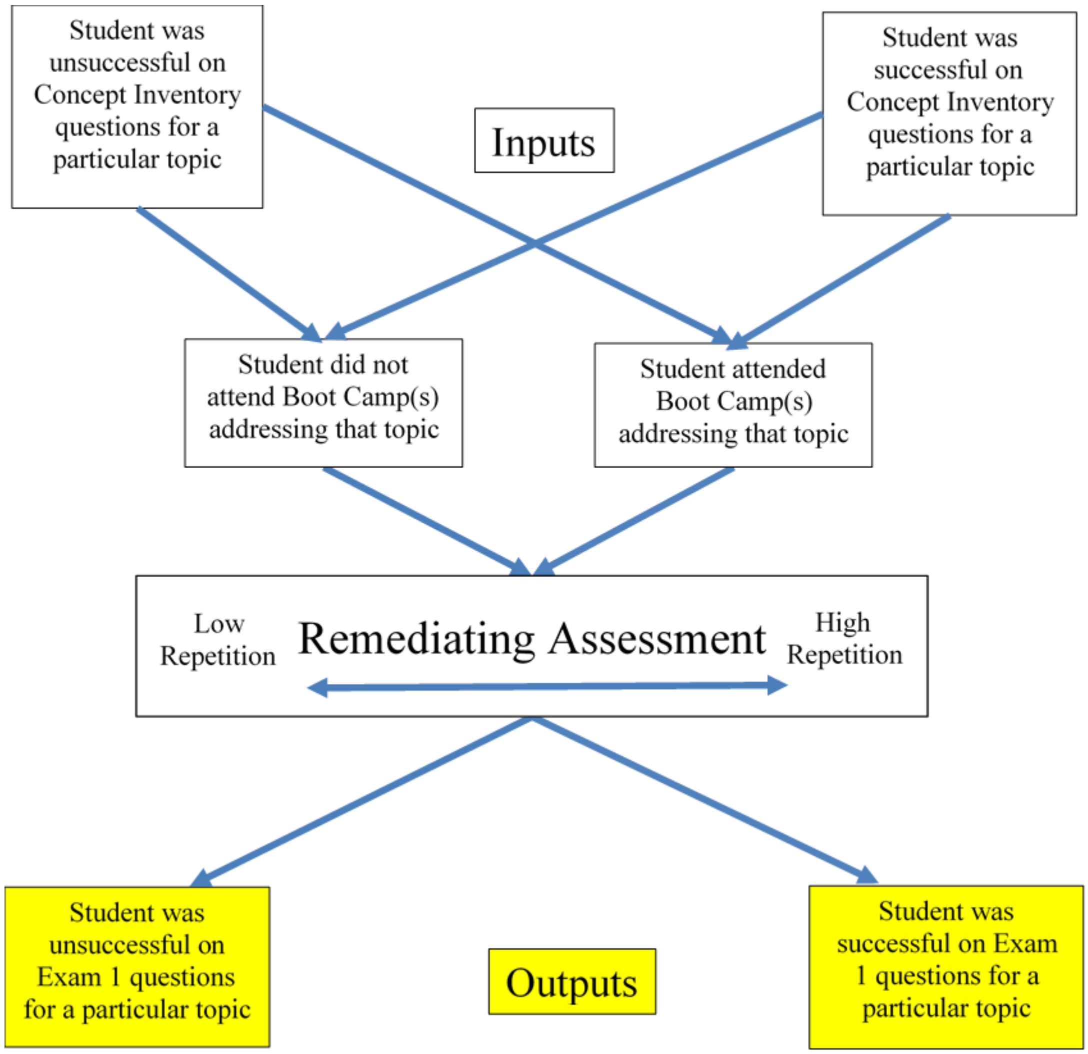
Statistics background of students were self-reported by the students as part of their initial Concept Inventory in the 2018 and 2019 course offerings. Students could report as having no statistics background or course before this course (coded as "No Stats"), having some statistics within another course in high school (coded as "Some HS"), having a dedicated statistics course in high school (coded as "HS"), having some statistics within another course in college (coded as "Some Col"), and having a dedicated statistics course in college (coded as "Col"). Students not completing the question were assigned Did Not Complete (coded as "DNC"). The 2016 cohort did not complete the LSSCI, thus student-self-reported statistical background is unknown. Student statistics background for the 2016 cohort was modeled from 2016 student major and student-reported statistical backgrounds from the 2018 and 2019 cohorts.

Path Model for Assessing effect of Course Structures

A path model framework proposed here was utilized to analyze how student Exam 1 outcomes were affected by three factors: (1) student pre-requisite knowledge as demonstrated by responses on Lean Six Sigma Concept Inventory (LSSCI), (2) student attendance at MLA-led statistics boot camps, and (3) number of student attempts at remediating assessments. Number of attempts on the remediating assessments was selected over final score, as it would be possible with enough attempts to view all questions with their answers, thus attaining a high or perfect score through error checking. The proposed path model tracked whether or not students engaged with these materials in the LSSCI, remediating assessments, and statistics boot camps, and checked if this engagement had a statistically significant effect on student outcomes documented in Exam 1 scores. Questions on the LSSCI at the beginning of the semester presented selected concepts in statistics. Guided and individual practice at the MLA-led statistics boot camps and remediating assessment questions as course homework presented these concepts as well. Finally, Exam 1 assessed these concepts summatively. A diagram of different paths students could take through the different iterations is displayed in **Figure 1Error! Reference source not found.**.

Figure 1

Path Model used for General Linear Modeling



The general linear model protocol begins with all possible contributing factor's data in each path listed as factors, with one of the Exam 1 question results listed as an output. The model then attempted to develop an equation that tried to predict output values based on a sum of linear functions of the factors, in the general form M_aX_a+M_bX_b+M_cX_c+M_dX_d… After the model calculation was completed by Minitab, if not all factors were significant (p-value <0.1), the least significant term was removed. This was repeated with fewer and fewer factors until: all factors were significant, or one non-significant factor remained. The first case provides a model that most efficiently explains the most variation in the output factors coming from variation in the remaining input factors. These results include the percentage of variation that can be attributed to the factors (Contribution), and variation that could not be attributed to any input factor tested (Error).

The second case with one remaining non-significant factor determines that none of the contributing input data factors contributed to variation in the output factor to a statistically significant degree.

The generalized path model was utilized to construct five path models for six key statistics concepts – central tendency, central tendency and dispersion synthesis, dispersion, union and intersection probability, z-chart comprehension, and accuracy versus precision. Using the General Linear Model framework in Minitab 19 (Minitab, 2019), analysis of the different amounts of students taking each individual path during the 2018 and 2019 courses was conducted to determine effects of the three factors – pre-requisite knowledge, attendance at statistics boot camps, and number of attempts on remediating assessments – had on successful student outcomes.

Path Analysis

Path Analysis consisted of utilizing Minitab's General Linear Model (GLM) to determine statistical significance of student learning outcomes based upon three factors: (1) scores on relevant concept inventory questions (Coded as CI 1 to CI 22); (2) attempts counts for each remediating assessment homework (Coded as AO1, AO2, AO3, AO4, and AO6), and (3) attendance in relevant boot camps listed in the path (Coded as BC 1 to BC 7). The GLM utilized main effects and second order interactions but excluded tertiary and high-order effects.

Considered outputs for the Path Analysis included one or more than one selected Exam 1 question based upon the path topic (Coded as EX 1, followed by the question number). After the GLM model was run, the main effect or second order interaction having the greatest *p*-value was removed from the factors list until all main effects and second order interactions had *p*-values less than $\alpha = 0.10$. A Path Analysis was completed for each of the five key statistical concepts. Remaining significant inputs were listed with contribution, F and *p*-value. If only one non-significant factor remained, no significant factors in the model were found to influence variation in the Exam 1 responses in a statistically significant manner.

Data Pruning

Two sets of data were removed from the Path Analysis. The first removed data set involved complete removal (*e.g.*, MLA-led boot camp, remediating assessments, Exam 1 responses) of data generated by two students. The two students were the only English as a Second Language (ESL) students in the study and submitted a large number of remediating assignment responses compared to their peers. Although unknowable, it is hypothesized that both students completed a large number of remediating assignments to develop and review vocabulary, not necessarily statistical topics. While removed from this analysis, the large number of remediating assessments by the ESL students suggests an additional benefit, *i.e.*, the ability of ESL students to remain with their cohort. Responses from both ESL students were excluded from the Path Analysis study.

The second data pruning was to eliminate any remediating assessment attempts having durations less than two minutes. Remediating assessments typically contained more than 15 questions, thus remediating assessment attempts having durations less than two minutes were assumed to be invalid attempts. They may have also been due to internet connectivity difficulties or difficulties with the Blackboard online software. Students also may have decided to not finish an attempt after starting it, submitting the attempt to close the window.

Remediating Assessment 5 (AO5) was not studied as a contributing factor in path analysis. This assessment was an assignment for students to complete the university's Information Security and Awareness training, before completing course projects. This was determined not to be a factor for study, as it a binary assignment (completed or not completed), that was not submitted more than once.

CHAPTER 4: RESULTS

Overall Examination Results

Exam 1 results by cohort are shown in **Error! Reference source not found.**. In **Error! Reference source not found.**, **N** indicates the number of students submitting Exam 1, $\boldsymbol{\bar{x}}$ is the algebraic mean, $\boldsymbol{s}$ is the sample standard deviation, and $\boldsymbol{z_\alpha}$ represents the 95% confidence interval defined as defined below.

$$\bar{x} = \frac{(\sum x)}{N} \qquad s = \sqrt{\frac{(x-\bar{x})^2}{N-1}} \qquad z_\alpha = z_{crit}\left(\frac{1-confidence\ level}{2}\right).$$

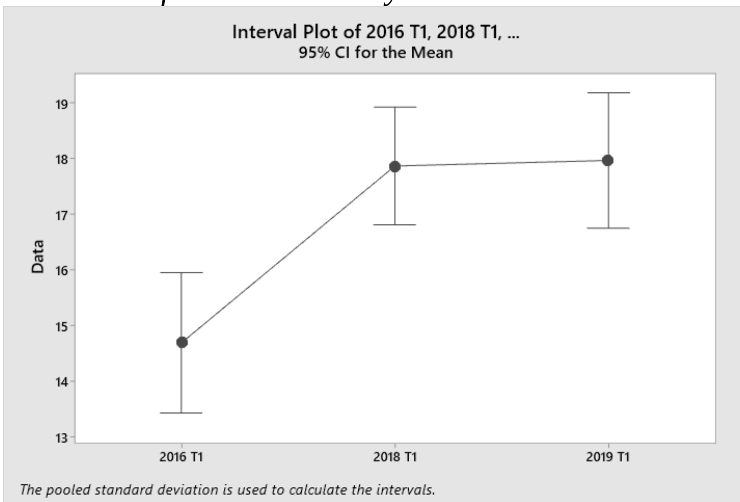
Table 3
Exam 1 results by cohort

Offering	N	Mean	St. Dev	95% CI
2016 T1	26	14.69	2.91	(13.43, 15.95)
2018 T1	37	17.87	3.16	(16.81, 18.92)
2019 T1	28	17.96	3.60	(16.75, 19.18)

Figure 2 graphically shows the data from **Error! Reference source not found.**. In this figure we see an approximately 3 point (out of 25) increase in average Exam 1 scores from 2016 to 2018, an increase that was maintained for the 2019 cohort.

Figure 2

Exam 1 Averages and 95% Confidence Interval by Cohort



A statistically significant difference between the three means was found using a one-way ANOVA, as demonstrated in **Table 4**. The difference was found to be between the 2016 and 2018 cohorts, and the 2016 and 2019 cohorts through subsequent two sample t-tests in Minitab 19. The mean scores of the 2018 and 2019 cohort were significantly higher than the 2016 mean score. The means of the 2018 and 2019 cohorts were not found to be statistically different from each other.

Table 4

Exam 1 Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	2	192.2	96.08	9.20	0.000
Error	88	918.8	10.44		
Total	90	1111.0			

These changes in the mean carried on throughout the semester, as seen **Error! Reference source not found.** for Exam 2, and **Error! Reference source not found.** for the Final Examination. Figure 3 shows that mean exam 2 scores improved approximately 7 points (out of

25) from the 2016 cohort to the 2018 cohort, and this improvement was maintained for the 2019 cohort.

Figure 3

Exam 2 Averages and 95% Confidence Interval by Cohort

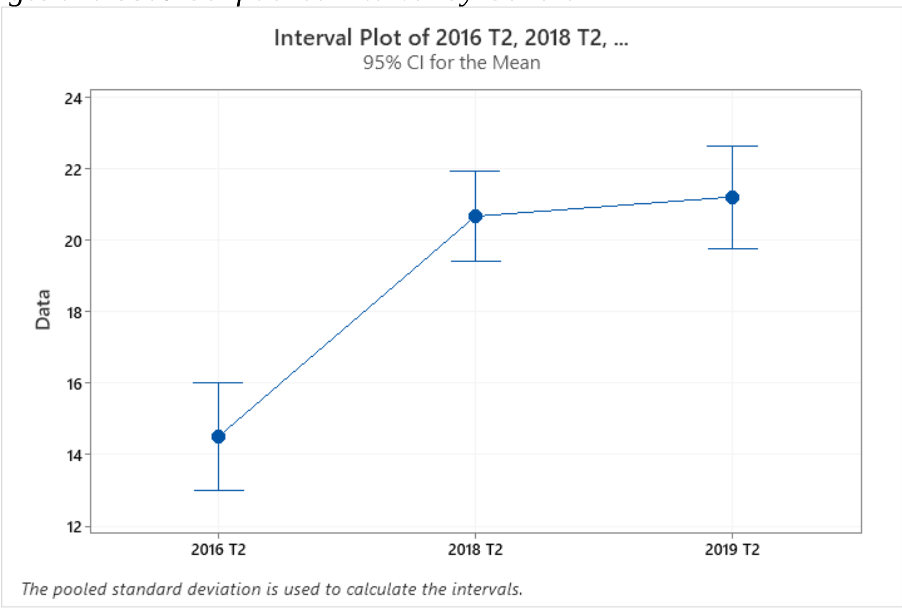
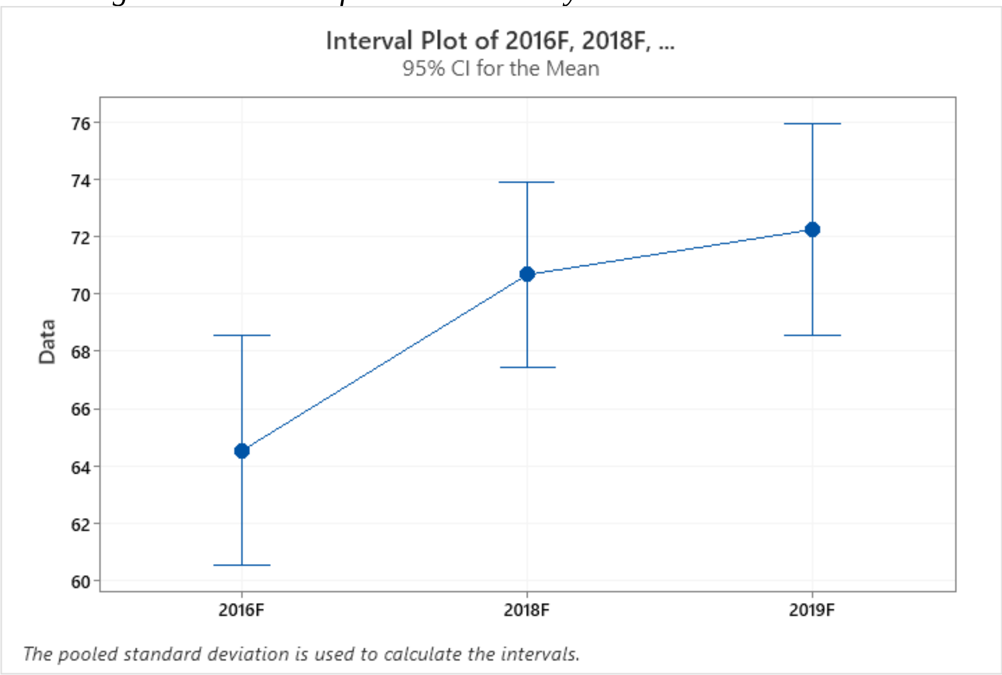


Figure 4 shows that mean final exam scores improved approximately 7 points (out of 100) from the 2016 cohort to the 2018 cohort, and this improvement was maintained for the 2019 cohort.

Figure 4
Final Exam Averages and 95% Confidence Interval by Cohort



In these cases, as shown in **Error! Not a valid bookmark self-reference.** and **Table 6Error! Reference source not found.**, the one-way ANOVA detected a difference in means, which was later found to be between the 2016 and 2018 cohorts, and the 2016 and 2019 cohorts. As with Exam 1 this was found with use of two-sample t-tests.

Table 5
Exam 2 Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	2	762.6	381.28	25.99	0.000
Error	88	1291.1	14.67		
Total	90	2053.6			

Table 6
Final Exam Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	2	852.9	426.45	4.37	0.016
Error	86	8391.3	97.57		

Total	88	9244.2

Examination of student statistical background by year

After determining the difference in outcomes for the identical Exam 1 from year to year, analysis of the student groups of the three course cohorts was conducted. This was performed to determine if there was a statistically significant difference in the students within each cohort that may have contributed to this change in exam outcomes. 2018 and 2019 student major and statistical background was collected from demographic information on Blackboard and through self-reporting on the course Concept Inventory. The reported majors and statistics backgrounds of students in the 2018 and 2019 cohorts (the two cohorts where this data was collected) are presented in **Table 7**.

Table 7

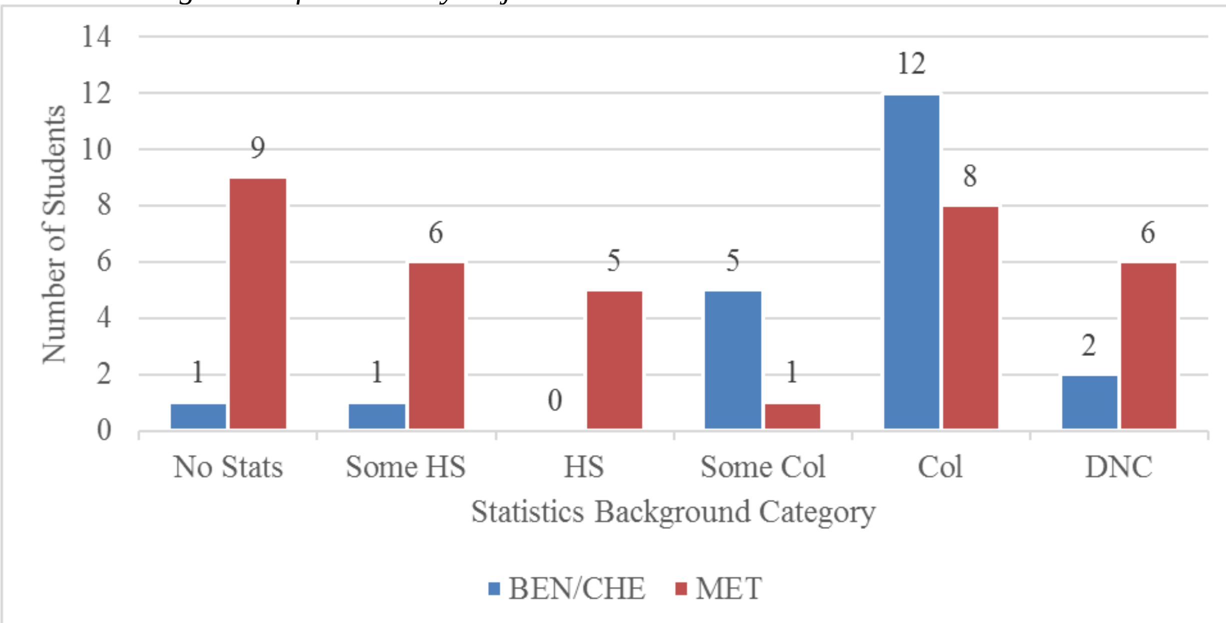
Statistics background of 2018 and 2019 cohorts, by major

Major	No Stats	Some HS	HS	Some Col	Col	DNC	Total
BEN/CHE	1	1	0	5	12	2	21
Grad	0	0	1	2	4	0	7
MEE	0	0	0	0	3	0	3
MET	9	6	5	1	8	6	35

These data show that the average background composition of students differs between those in chemical and bioengineering (CHB/BEN) and those in mechanical engineering technology (MET) programs. This composition especially differs in the percentages of each category of statistics background, as seen in below.

Figure 5

Statistics background of students by major



Since statistics backgrounds of the 2016 offering were not collected, but majors were, inferred estimates of the statistics background concentrations of the 2016 course offering were made. These were calculated using the major concentrations of the 2016 course offering, and the average statistic history percentages of these majors over the 2018 and 2019 course offerings. As an example, the 2016 cohort had a higher percentage of MET students than the other two, and METs were more likely than other groups to have a background of no statistics, so the estimated “No Stats” subgroup is larger than this subgroup in the two other cohorts. The estimates are listed below in number of students in Table 8, and as percentages of the class at large in Table 9 below.

Table 8

Estimated statistics background for 2016 cohort based on 2018 and 2019 backgrounds for MET and BEN/CHE students.

Major	No Stats	Some HS	HS	Some Col	Col	DNC	Total
BEN/CHE	0.24	0.24	0	1.19	2.85	0.476	5
MET	5.4	3.6	3	0.6	4.8	3.6	21

Overall	5.64	3.84	3	1.79	7.65	4.076	26

Table 9

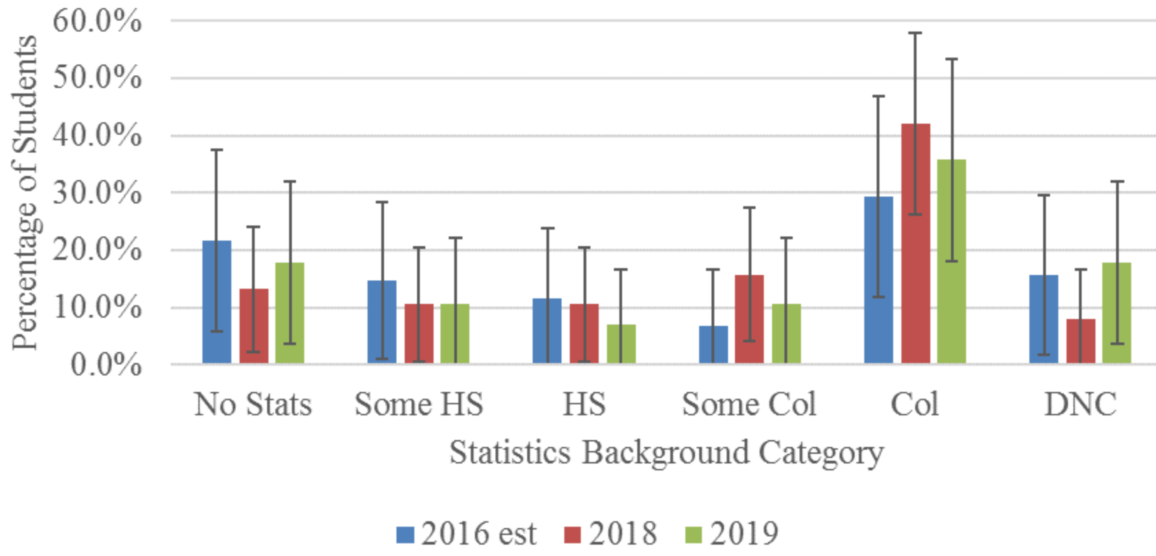
Statistics background percentages by cohort

Cohort	No Stats	Some HS	HS	Some Col	Col	DNC	Total
2016 estimate	21.7%	14.8%	11.5%	6.9%	29.4%	15.7%	100.0%
2018	13.2%	10.5%	10.5%	15.8%	42.1%	7.9%	100.0%
2019	17.9%	10.7%	7.1%	10.7%	35.7%	17.9%	100.0%

Data from Table 9 are shown in

Figure 6, in graphical format with 95% confidence interval expressed in the error bars. The sample sizes of the three course offerings (N = 26, N = 38, N = 30, in chronological order) ensure that these margins of error are quite large in relation to the calculated proportion. No statistically significant differences in proportion were found for any statistics background group on a year by year basis were found using several ANOVAs. This and the fact that none of the majors involved altered their requirements for college statistics over the time interval studied, suggests that there is not a difference in the statistics background of the 2016 group that would possibly create the significant difference in student outcomes.

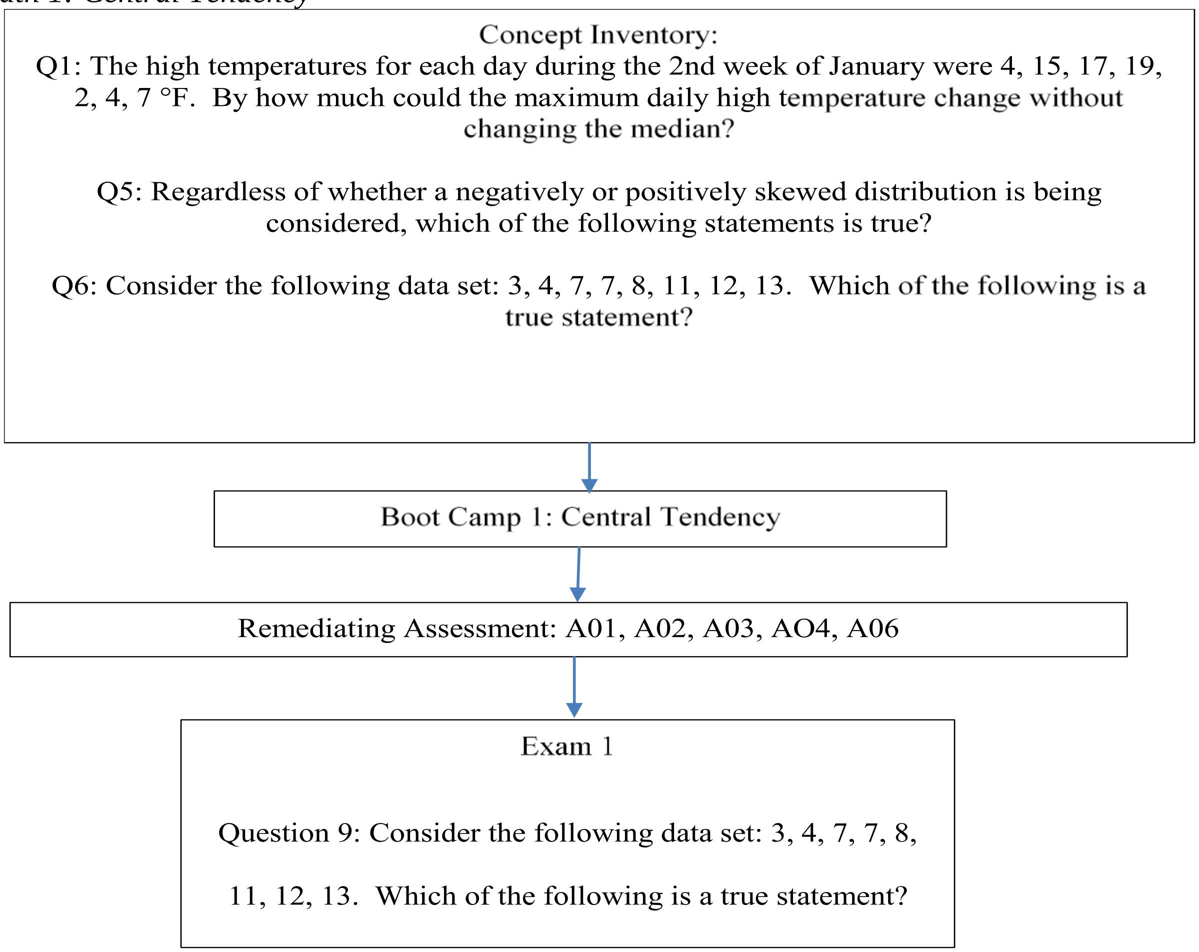
Figure 6
Statistics background percentage by year (with 95% confidence interval)



Path 1 Analysis: Central Tendency

Path 1 analyzed central tendency (*e.g.*, mean, median and mode) throughout the concept inventory questions on the topic, the boot camp focusing on this topic (Boot Camp #1), the remediating assessments (AO1, AO2, AO3, AO4, and AO6), and finally through the output of an Exam 1 question, where students were asked to report the relationship between the mean and median of the data set. The Exam 1 question, Question 9, required students to calculate both values using their definitions of both. Detailed wording of questions is present on the path diagram below in **Error! Reference source not found.**.

Figure 7
Path 1: Central Tendency

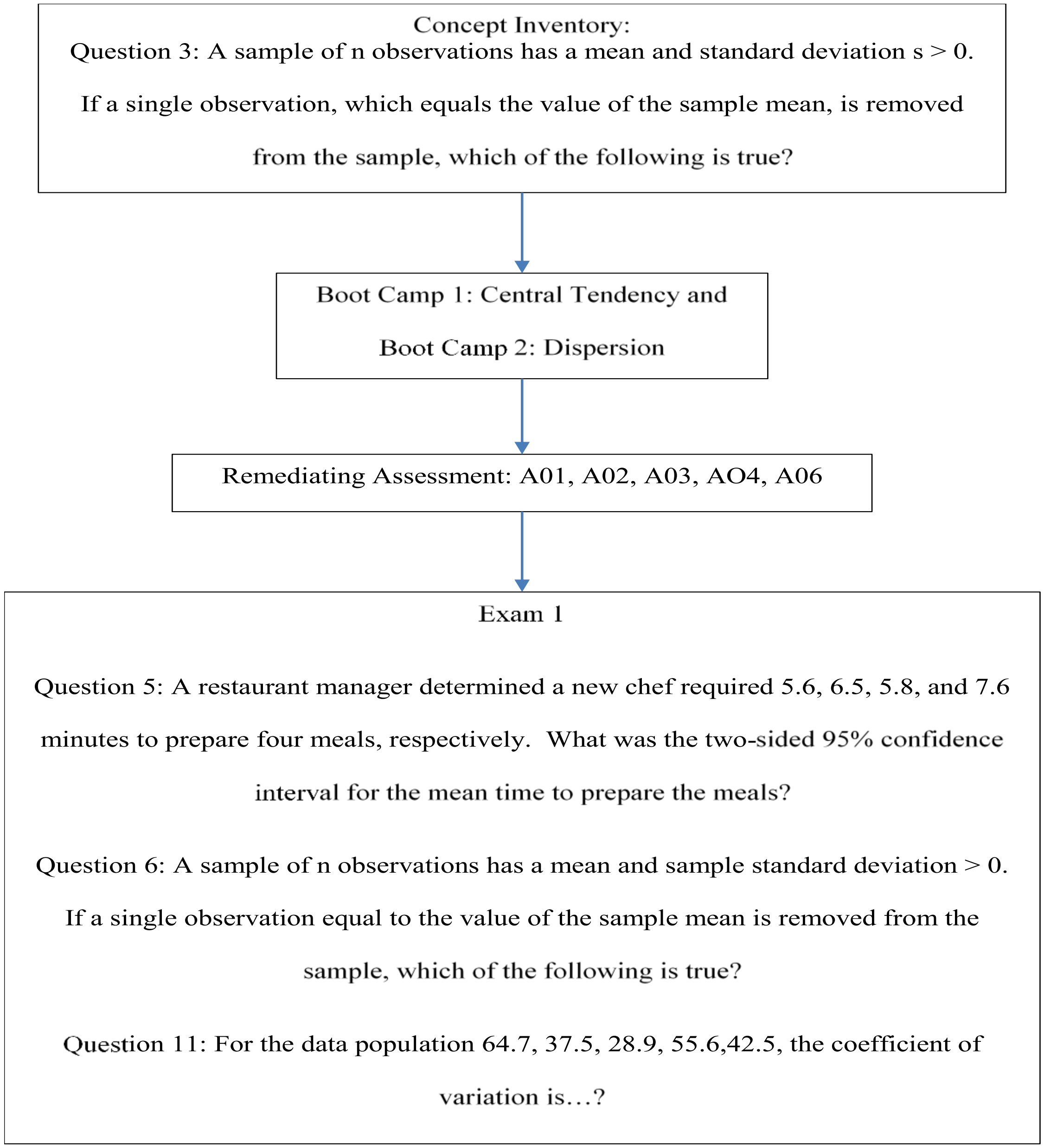


The analysis of the GLM showed no significant input factors due to inability to accurately collect usable *p*-values for each factor. This arose from the fact that nearly every student answered these questions correctly both at the beginning of the path (the Concept Inventories) and the end of the path (on the Exam 1 questions). Thus, with all the paths having the same start and end point, useful data on the effect of the intermediate steps on the outcomes was not found in this situation.

Path 2 Analysis: Central Tendency and Dispersion synthesis

Path 2 analyzed the synthesis of concepts of central tendency measure (*e.g.*, mean, median and mode) and concepts relating to data dispersion (*e.g.*, standard deviation, variance). These concepts were tracked via concept inventory questions on the synthesis of these topics, the boot camps focusing on this topic (Boot Camp #1 and #2), the remediating assessments (AO1, AO2, AO3, AO4, and AO6), and finally through the output of Exam 1 questions. In these exam questions, students were asked to report on values that required them to calculate both measures of central tendency and dispersion, such as a confidence interval and a coefficient of variation. The Exam 1 questions also required students to understand the relationships between the two measures, especially in Exam 1 Question 6. Figure 8 presents detailed wording of questions on the path diagram below.

Figure 8
Path 2: Central Tendency and Dispersion



The following three tables, **Table 10Error! Reference source not found.**,

Table **11**

Path 2 Analysis of Variance (Exam 1 Q6) and

Table **12**, present the General Linear Model findings for each Exam 1 question. In all of the following results, the factors listed from the concept inventory, boot camps, and remediating assessments were found to have a statistically significant contribution to the variation of performance on the Exam 1 questions specified. This includes remediating assignments AO1, AO3, and AO6, and Concept Inventory question 3 for Exam 1 questions 5 and 6. Attendance in Boot Camp #1 and #2 covaried with Exam 1 question 11. The three linear models have R^2 values of 62.96%, 64.32%, and 17.34% respectively, showing moderate to slight explanation of variance in the output variable's value, by these input variables.

Table 10

Path 2 Analysis of Variance (Exam 1 Q5)

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
AO1	8	33.96	15.34%	47.12	5.890	2.23	0.053
AO3	9	63.58	28.71%	52.94	5.883	2.22	0.048
AO6*CI 3	7	41.86	18.91%	41.86	5.981	2.26	0.056
Error	31	82.03	37.04%	82.03	2.646		
Total	55	221.43	100.00%				

Table 11

Path 2 Analysis of Variance (Exam 1 Q6)

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
AO1*CI 3	8	43.84	23.14%	38.23	4.778	2.19	0.056
AO3*CI 3	9	32.67	17.25%	41.81	4.646	2.13	0.057
AO6*CI 3	7	45.33	23.93%	45.33	6.475	2.97	0.017
Error	31	67.59	35.68%	67.59	2.180		
Total	55	189.43	100.00%				

Table 12

Path 2 Analysis of Variance (Exam 1 Q11)

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
BC 1	1	26.08	12.02%	22.53	22.527	6.66	0.013
BC 1*BC 2	1	11.53	5.32%	11.53	11.531	3.41	0.070
Error	53	179.25	82.66%	179.25	3.382		
Total	55	216.86	100.00%				

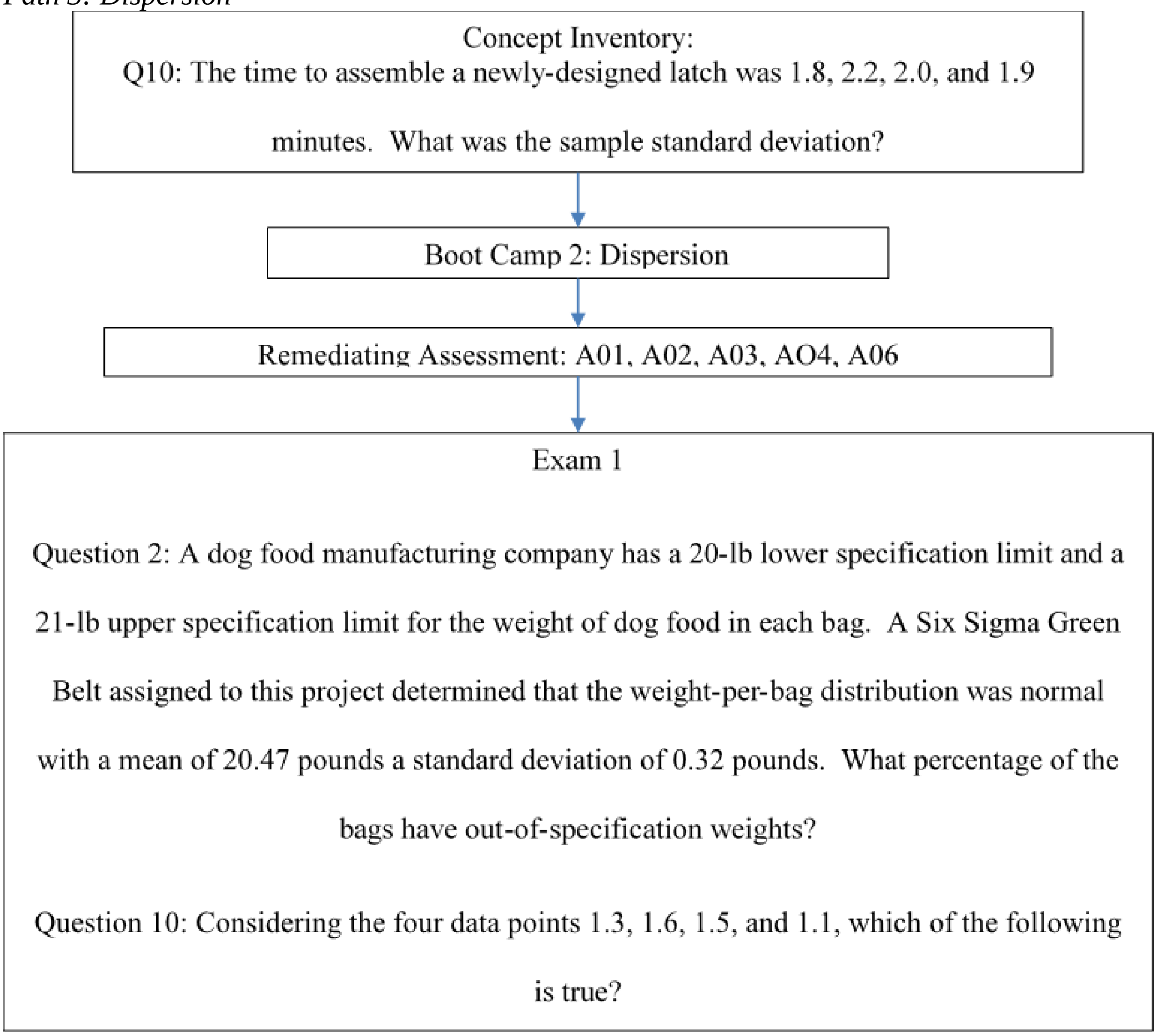
Path 3 Analysis: Dispersion

The third path analyzed the concept of data dispersion (*e.g.*, standard deviation) throughout the concept inventory questions on the topic, the boot camp focusing on this topic

(Boot Camp #1), the remediating assessments (AO1, AO2, AO3, AO4, and AO6), and finally through the output of an Exam 1 question, where students were asked to calculate the standard deviation of the data set. This was either done as an intermediate step in a multi-step problem (Exam 1 Question 2) or was the object of the problem itself (Exam 1 Question 10). Detailed wording of questions is present on the path diagram below in

Figure **9**.

Figure 9:
Path 3: Dispersion



The following tables, Table 13 and Table 14, present the General Linear Model findings for each Exam 1 question. As before, factors that covaried with Exam 1 question performance for the Exam 1 questions in this path are listed in each table. Remediating assignment AO3 impacted Exam 1 question 2, and Boot Camp #2 attendance covaried with Exam 1 question 10. The two linear models have R^2 values of 19.41%, and 9.92% respectively, showing slight explanation of variance of the output factors by these input factors.

Table 13:
Path 3 Analysis of Variance (Exam 1 Q2)

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
AO3	9	46.83	19.41%	46.83	5.204	1.42	0.204
Error	53	194.44	80.59%	194.44	3.669		
Total	62	241.27	100.00%				

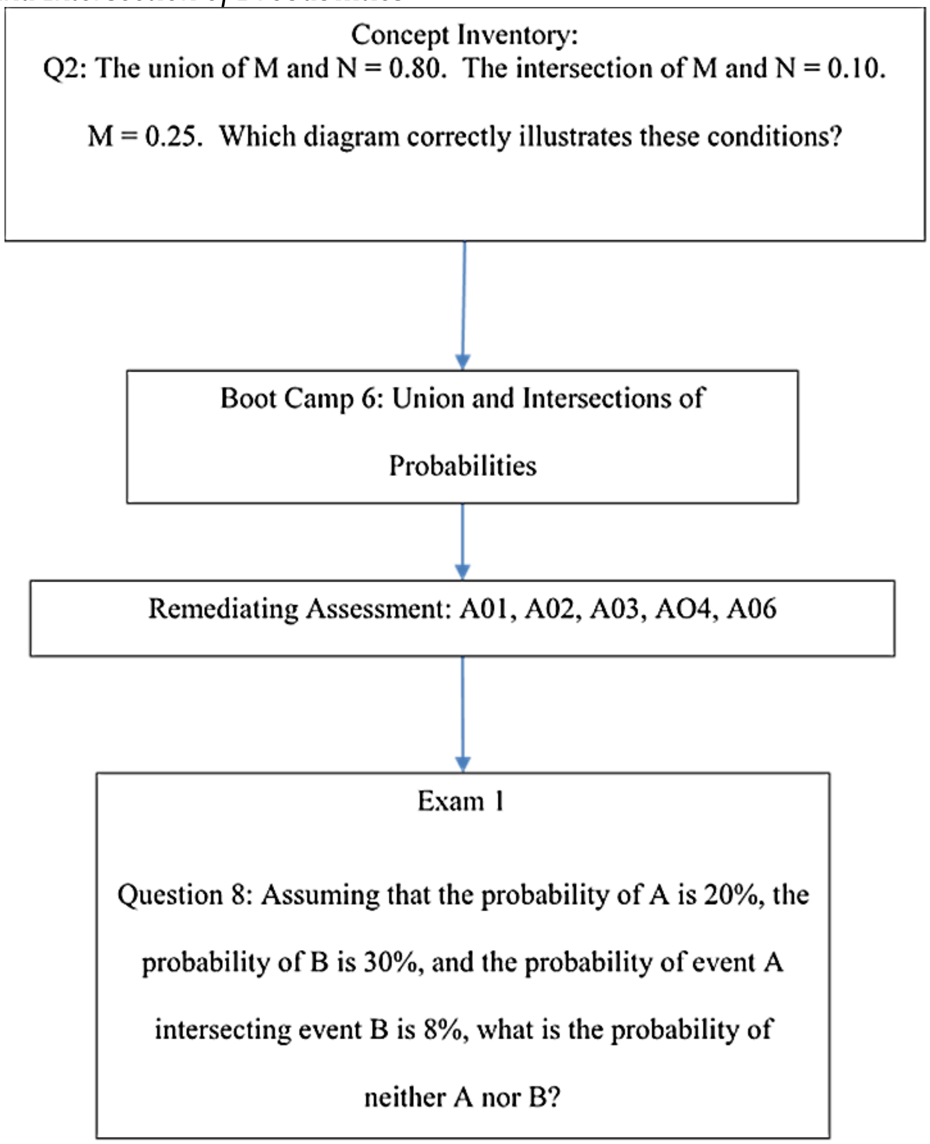
Table 14:
Path 3 Analysis of Variance (Exam 1 Q10)

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
BC 2	1	18.95	9.92%	18.95	18.949	6.72	0.012
Error	61	172.03	90.08%	172.03	2.820		
Total	62	190.98	100.00%				

Path 4 Analysis: Union and Intersection of Probability

Path 4 analyzed the concept of probability, in particular in how it pertained to the union and intersection probabilities of two more non-mutually exclusive events. This was tracked through throughout the concept inventory questions on the topic, the boot camp focusing on this topic (Boot Camp #6), the remediating assessments (AO1, AO2, AO3, AO4, and AO6), and finally through the output of an Exam 1 question, where students were asked to calculate the probability of two events happening based on the given probability of their union and intersection. Detailed wording of questions is present on the path diagram below in Figure 10.

Figure 10:
Path 4: Union and Intersection of Probabilities



General Linear Model findings found no significant input factors present that covaried with the output of this path, Exam 1 Question 8. The remaining variable after linear model analysis had a *p*-value exceeding our alpha value of 0.1.

Path 5 Analysis: Z-chart comprehension

The fifth path analyzed the concept of reading Z-charts, to find probability of normally distributed probabilities. This was modeled using the concept inventory questions on the topic, the boot camp focusing on this topic (Boot Camp #3), the remediating assessments (AO1, AO2, AO3, AO4, and AO6), and finally through the output of an Exam 1 question, where students were asked to calculate the z-value of a data point in a set given a population mean and standard deviation. Students then needed to apply this z-value to determine what percentage of normally distributed data points would be more extreme than the data points given. Detailed wording of questions is present on the path diagram below in **Figure 11**.

Figure 11

Path 5: Z-charts

Concept Inventory:

Q11: Pat's Pizza determined that the amount of cheese on a pizza is normally distributed having a mean = 75.2 grams and a standard deviation = 3.4 grams. A pizza was measured to have 76.3 grams of cheese. What is the z value?

Q12: Assuming a normal distribution, what is the probability of an event with a response greater than a z-value of 1.37?

Q13: Assuming a normal distribution, what is the probability of an event with a response less than a z-value of 0.58?

Boot Camp 3: Z-Charts

Remediating Assessment: A01, A02, A03, AO4, A06

Exam 1

Question 2: A dog food manufacturing company has a 20-lb lower specification limit and a 21-lb upper specification limit for the weight of dog food in each bag. A Six Sigma Green Belt assigned to this project determined that the weight-per-bag distribution was normal with a mean of 20.47 pounds a standard deviation of 0.32 pounds. What percentage of the bags have out-of-specification weights?

Question 3: A dog food manufacturing company has a 20-lb lower specification limit and a 21-lb upper specification limit for the weight of dog food in each bag. A Six Sigma Green Belt assigned to this project determined that the weight-per-bag distribution was normal with a mean of 20.47 pounds a standard deviation of 0.32 pounds. Two projects are being considered to reduce the percentage of incorrectly filled bags. The first project will center the process within the specification limits. The second project will reduce the standard deviation to 0.30 pounds. Both projects cost $1,000. Which project do you recommend?

Table 15 and **Table 16** present the General Linear Model findings for each Exam 1 question. Successful completion of Concept Inventory question 13 covaried with Exam 1 question 2. Remediating assessment AO3 and successful completion of Concept Inventory question 12 covaried with Exam 1 question 13. The linear models have R^2 values of 6.98%, and 32.21% respectively, showing slight explanation of variance by these input variables.

Table 15
Path 5 Analysis of Variance (Exam 1 Q2)

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
CI 13	1	14.02	6.98%	14.02	14.018	4.05	0.049
Error	54	186.84	93.02%	186.84	3.460		
Total	55	200.86	100.00%				

Table 16
Path 5 Analysis of Variance (Exam 1 Q3)

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
AO3	9	40.05	22.79%	49.68	5.520	2.09	0.051
CI 12	1	16.54	9.41%	16.54	16.540	6.25	0.016
Error	45	119.12	67.79%	119.12	2.647		
Total	55	175.71	100.00%				

Path 6 Analysis: Accuracy versus Precision (Control)

The sixth and final path analyzed the concept of accuracy and precision, in particular how data sets are categorized using the two descriptors, which have different statistical meanings. This was tracked through the concept inventory question on the topic, the remediating assessments (AO1, AO2, AO3, AO4, and AO6), and finally through the output of an Exam 1 question, where students were asked to describe data points on a target diagram as accurate, precise, both or neither. There was no boot camp instruction on this topic, this analysis was conducted to investigate the baseline changes in paths students would take without boot camp instruction. Detailed wording of questions is presented on the path diagram below in

Figure **12**.

Figure 12:
Path 6: Accuracy vs. Precision

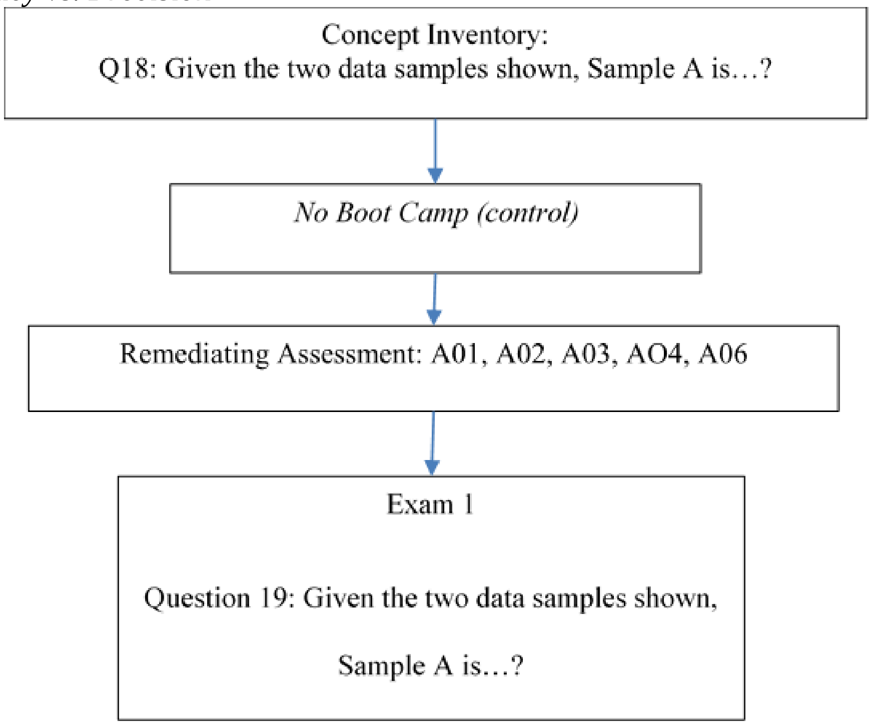


Table 17 presents the General Linear Model findings for the Exam 1 question. Successful completion of Concept Inventory question 18 covaried with Exam 1 question 19. The linear model has a R^2 value of 9.19%, showing slight explanation of variance by these input variables. This explanation was expected to be larger, as in this case the concept inventory question and examination question were identical.

Table 17:
Path 6 Analysis of Variance (Exam 1 Q19-Control)

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
CI 18	1	4.176	9.19%	4.176	4.1764	5.47	0.023
Error	54	41.252	90.81%	41.252	0.7639		
Total	55	45.429	100.00%				

What changes are seen in the three course offerings studied, regarding outcomes of demonstrated statistics knowledge and skills?

Significant improvement in Exam 1 success was observed from Fall 2016 to Spring 2018. This improvement in Exam 1 performance was carried forward into the Spring 2019 course offering. Through use of a one-way ANOVA on examination scores from the three course offerings, statistically significant improvement from Fall 2016 to Spring 2018 was seen also on Exam 2, and the comprehensive final exam. In both cases, this marked improvement was maintained between the Spring 2018 and Spring 2019 course offerings, between which there was no statistically significant differences in examination scores for any exam studied.

What differences exist regarding the background knowledge of students regarding statistics in the three course offerings studied?

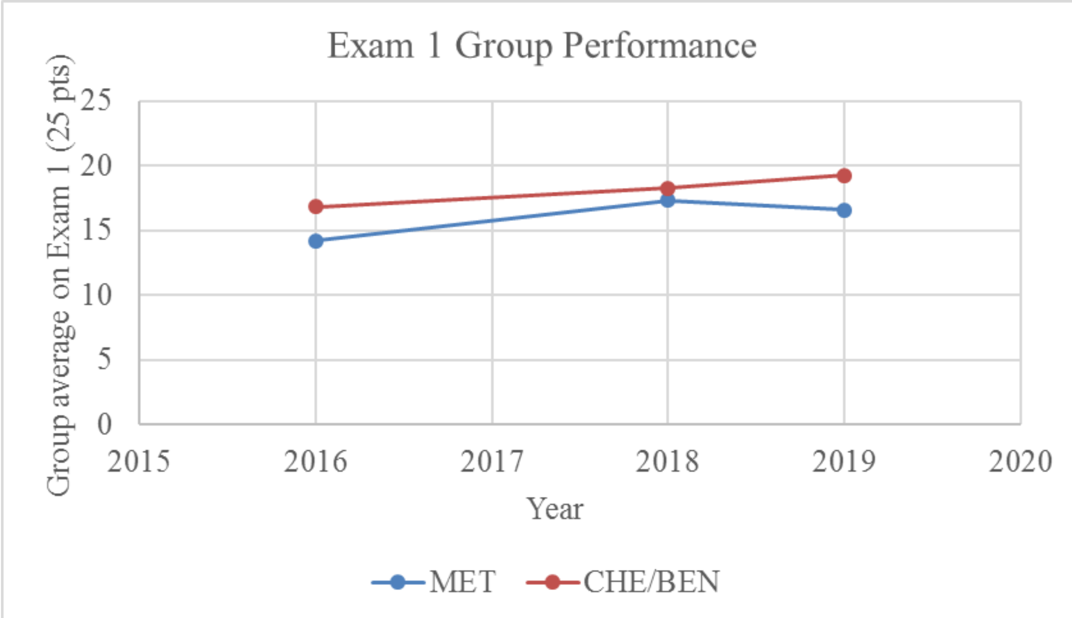
There was no statistically significant difference in the composition of the two course cohorts where background knowledge was recorded, in regard to reported statistics background. Inferences were made regarding the composition of the Fall 2016 cohort based on recorded data for the enrolled student majors. While there was a difference in average statistics background of MET students and CHB/BEN students, the class composition percentages of these two groups did not show a statistically significant change over the three course offerings studied. Thus, the percentages overall of students in each statistics background category did not change significantly over the course offerings. This combined with none of the majors in question having changed their curriculum or requirements with regards to statistics, suggests that the three course offerings had approximately equal statistics background composition.

What contributions did adding the course structures of statistics boot camps and remediating assessment have on student ability to demonstrate statistics knowledge in the 2018 and 2019 course offerings?

Weak to no direct change in Exam 1 success was seen by path analysis in connection to intermediate step completion, such as concept inventory question successful completion, Boot camp attendance and remediating assessment attempts. This lack of direct influence was most surprising when the effect of concept inventory question success is considered: one could reasonably assume that students completing a question correctly would have a better chance of being successful on the same or similar problem 7 weeks later.

Looking at the change of performance on Exam 1 for our two main study groups (CHE/BEN and MET), as we see in Figure 13, a performance gap between the two in the 2016 cohort, that shrinks in the 2018 cohort. The gap widens again in 2019, but both groups averages are still significantly above their 2016 levels.

Figure 13:
Exam 1 average by year for studied student groups



The current small sample size of student data available for study (75-80 students) limits resolution to be able to see smaller gains and changes accurately through quantitative methods. This size was one motivation for moving the general linear model framework used from testing multiple outputs (Exam 1 scores) at once, to testing each output individually against all possible inputs. This limiting of degrees of freedom required more tests to be run, but allowed for the most possible resolution. However even with this choice made, effect size seen of any intermediate step would have had to be very large and very consistent across 2018 and 2019 student cohorts in order to determine a strong statistical correlation of output effect.

CHAPTER 5: CONCLUSIONS AND LIMITATIONS

Although student outcomes in the Spring 2018 and Spring 2019 offerings were significantly improved over the student outcomes in the Fall 2016 offering, the limited sample size precludes assigning single factor causality for the improvement. Our analysis of the student's background in statistics before entering the course did not show a difference in actual and/or estimated composition between the three cohorts, that would have likely skewed the data. No one intervention was solely responsible for the variation seen in Exam 1 performance, even though a significant variation was shown in marked improvement after the interventions were implemented. This trend was even seen when it would be reasonable to expect a strong source of variation. In several cases, the concept inventory questions and examination questions were very similar or even identical. The concept inventory question success factor, however, was never found to be a factor explaining more than 30% of the output factor variation. This finding suggests that the relationship between student's performance on a question and their performance on a very similar or identical question seven weeks later was more complicated than could be explained by a simple relationship of causation. This simple argument would be that if they were successful at the beginning, they should be successful at the end. However, this was not found to be the case.

This finding also leads to the conclusion that all the interventions together contributed more towards student success increases as a whole, more so than any factor influenced them separately. This conclusion is illustrated in the differences between a possible model in

Figure 14

Contributing Factors Model

Figure 14, and a more accurate model consistent with the findings illustrated in Figure 15. The more accurate model does this by the use of multiple arrows demonstrating the diverse contributing factors leading to student success, compared to one big arrow. In this case, we do not get a quick and simple answer for what we need to do to have students instantaneously be guaranteed increased success in a course. This makes sense, since student background knowledge is diverse and complex, as is how each student learns. Some students may have benefitted more from one intervention or another, and some may have simply benefitted from there being an intervention at all. All interventions involved the students having an opportunity to interact with course material in a setting and context they would not have been able to otherwise. By increasing the amount of time and repetitions students saw and worked with statistical concepts, students increased their retention of information and fluency with the tools of statistical inquiry. The proverbial rising tide of the interventions lifted all boats, even if we can't pin exactly where the water is coming from.

Figure 14

Contributing Factors Model 1- Possible Model

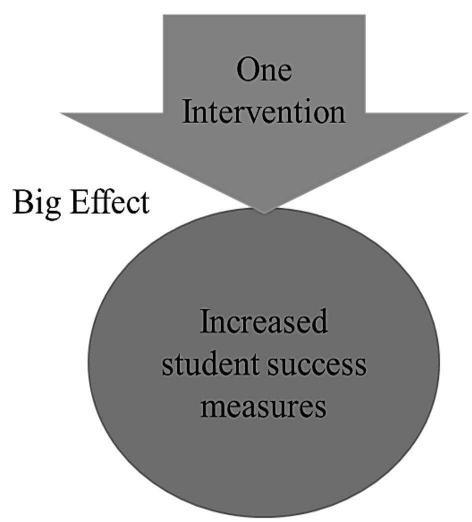
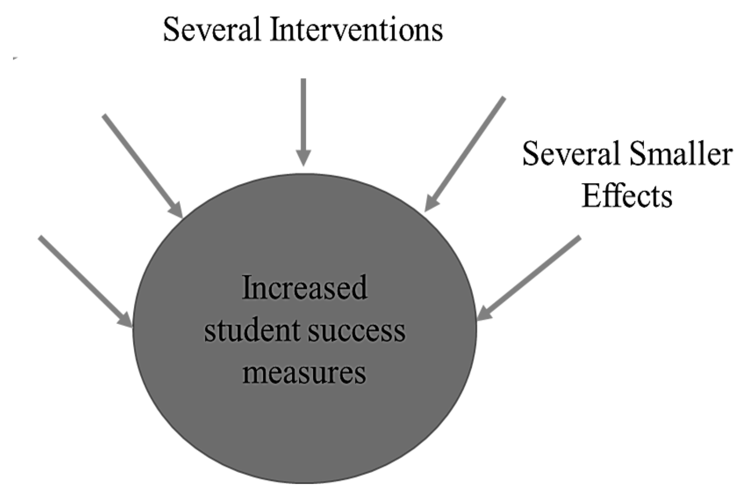


Figure 15
Contributing Factors Model 2- Supported Model



This also is supported by the previously mentioned results from analyzing the effects on student effects of similar interventions separately. Howard's (2019) study of remediating assessment and Liu's (2017) study in adaptive assessment both found that these interventions had either small or no significant effect on outcomes for students, and any effects seen were only seen in students with small gaps in background knowledge. Using these results combined with the conclusions of this study, one can hypothesize that these interventions taken on their own may not increase student engagement with material enough to produce measurable change in student outcomes later on. However, as part of a group of several interventions, each allowing students to work with the material in different environments and ways, the interventions may work synergistically to improve course outcomes more than any one intervention could do alone.

Developing a system of several interventions for students with background knowledge gaps may help to increase engagement and effort of students where these gaps are substantial. The use of a single method for intervention may cause students to become discouraged if that particular method does not work well or promptly for them in addressing their knowledge gaps. These gaps may then appear insurmountable for the student. Howard (2019) also supports this with their data, where they found that remediating assessment alone benefitted for the most part only students where their background knowledge gaps were small, while students with large knowledge gaps and low achievement were less affected by the intervention. However, in our data with combined interventions, both student groups studied (CHE/BEN and MET) were shown to have improved, despite both groups having different levels of background knowledge.

The very act of offering several different modes of intervention regularly throughout the course may help students enforce skills for looking at material outside of assignments and studying independently. It also shows students that the instructor is invested in addressing their background knowledge gaps and has put effort into a systemic plan to help the student improve. This can help turn “I don’t have the necessary background knowledge” to “I don’t have the necessary background knowledge *yet*” in students’ minds.

This method of intervention may be especially helpful for so-called “barrier” courses in STEM disciplines, where some students may require more support in background knowledge ideas in order to succeed in the courses (Baker-Ward et al., 1993; Budny et al., 2002). The interventions studied in this paper as a system were found to positively affect student success outcomes. This encourages the use of similar programs in courses and curriculums where more support is required. This could increase student retention in STEM disciplines in general, not just engineering. Students that could benefit most from this are students that are currently underrepresented minorities in STEM, who are more likely than average to require more background knowledge support, especially in earlier course due to decreased access to support resources in earlier education (National Committee on Science, Engineering, 2011)

Our results showed a positive effect on student success from a set of interventions including MLA-led statistics boot camps. This reinforces previous research on the effectiveness of student-driven instruction and active learning framework for learning. Maine Learning Assistants were able through the boot camps to connect the background knowledge concepts with Lean Six Sigma concepts being addressed in class, helping to increase the perceived value of the concepts to the students. This was found to generate positive effects by Harackiewicz (2018) and Acee (2010) as well, which together with our results encourages continued MLA interaction with this course.

Regarding recommendations for course instructors of this particular course, this researcher would encourage the development of several interventions that are aimed at helping remediate student background knowledge. These interventions should allow different pathways for students to interact with fundamental course concepts, and should be developed for review and retention over the time of the course. These could be the ones analyzed in this study or could change based on instructor or MLA observation of which supports help students with particular concepts. While we did see a rise in student outcomes after the interventions, the data collected does not tie this result back to any particular intervention. Thus, developing a system that best addresses these knowledge gaps may require further adaptation or experimentation. New interventions can be added, and some could be discontinued, with student data and feedback influencing those decisions. In this case we might not be able to quantitatively prove that an intervention is influencing students, but we can ask the students what is helping them.

These results and conclusions do not necessarily deem quantitative data collection for this type of intervention impossible, but in this case the data volume required may not be reasonable or prudent. A study completed here at the University of Maine on an introductory biology course with similar interventions was able to find statistically significant improvement with a larger data set of 700 students (Batz et al., 2015), as far more students have taken Biology 100 in the last three years than have taken MET 440. This improvement was able to be directly tied to one particular intervention, learning assistant delivering tutoring sessions after a screening assignment. Estimating from average MET 440 enrollment numbers, to reach this sample size would take approximately 22 course offerings, which if we include the previous offerings would still require 38 additional years of identical instruction and data collection. Bearing this in mind, further data collection should consider the reasonability of data collection, and therefore should move towards more qualitative modes of data collection and analysis. These methods may be more suited for the types of contributions present in this course, and more compatible with our available sample size.

An example of possible qualitative data to collect is to characterize MLA pedagogy before, during, and after teaching Statistics boot camps. Preliminary testimony from the Maine Learning assistant in Spring 2018 detailed that the experience of delivering this intervention in a MET 440 class changed their perceptions of how teaching and learning at the collegiate level should be valued and can be improved, eventually leading the learning assistant to shift focus of their graduate study towards STEM education. Further clinical interviews of other learning assistants for the course and perhaps group interviews or discussions of students who attended the boot camps and used other interventions, may be informative in how the interventions qualitatively alter student and teacher understanding of the course objectives.

BIBLIOGRAPHY

Acee, T. W., & Weinstein, C. E. (2010). Effects of a value-reappraisal intervention on statistics students' motivation and performance. *Journal of Experimental Education*, *78*(4), 487–512. https://doi.org/10.1080/00220970903352753

Ali, F., & Ahmed, A. (2016). A review on the statistical concepts of six sigma. *AIP Conference Proceedings*, *1775*. https://doi.org/10.1063/1.4965173

Baker-Ward, L., Mohr, P. H., Dietz, E. J., Forrest, K. D., & Felder, R. M. (1993). A Longitudinal Study of Engineering Student Performance and Retention: I. Success and Failure in the Introductory Course. *Journal of Engineering Education*, *82*(1), 15–21. https://doi.org/10.1002/j.2168-9830.1993.tb00067.x

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bateiha, S., Marchionda, H., & Autin, M. (2020). Teaching style and attitudes: a comparison of two collegiate introductory statistics classes. *Journal of Statistics Education*, *0*(0), 1–16. https://doi.org/10.1080/10691898.2020.1765710

Batz, Z., Olsen, B. J., Dumont, J., Dastoor, F., & Smith, M. K. (2015). Helping struggling students in introductory biology: A peer-tutoring approach that improves performance, perception, and retention. *CBE Life Sciences Education*, *14*(2), 1–12. https://doi.org/10.1187/cbe.14-08-0120

Bragdon, D. (2014). University students' graph interpretation and comprehension abilities. *Electronic Theses and Dissertations*, *2090*. https://doi.org/10.1080/19477503.2018.1480862

Budny, D., Bjedov, G., & LeBold, W. (2002). Assessment of the impact of the freshman engineering courses. *Journal of Engineering Education*, *87*(4), 1100–1106. https://doi.org/10.1109/fie.1997.636047

Daun-Barnett, N., & St. John, E. P. (2015). Constrained Curriculum in High Schools: The Changing Math Standards and Student Achievement, High School Graduation and College Continuation. *Education Policy Analysis Archives*, *20*, 5. https://doi.org/10.14507/epaa.v20n5.2012

Gal, I. (2004). Statistical Literacy: meanings, components, responsibilities. In *The challenge of developing statistical literacy, reasoning and thinking* (pp. 47–78).

Gal, Iddo, & Garfield, J. (1997). Curricular Goals and Assessment Challenges in Statistics Education. *The Assessment Challenge in Statistics Education*, *1997*, 1–13.

George, M. L., Rowlands, D., Price, M., & Maxey, J. (2005). *The Lean Six Sigma Pocket Toolbook*. *The Lean Six Sigma Pocket Toolbook*.

Harackiewicz, J. M., & Priniski, S. J. (2018). Improving Student Outcomes in Higher Education: The Science of Targeted Intervention. *Annual Review of Psychology*, *69*(1), 409–435. https://doi.org/10.1146/annurev-psych-122216-011725

Hieb, J. L., Lyle, K. B., Ralston, P. A. S., & Chariker, J. (2015). Predicting performance in a first engineering calculus course: implications for interventions. *International Journal of Mathematical Education in Science and Technology*, *46*(1), 40–55. https://doi.org/10.1080/0020739X.2014.936976

Howard, E., Meehan, M., & Parnell, A. (2019). Quantifying participation in, and the effectiveness of, remediating assessment in a university mathematics module. *Assessment and Evaluation in Higher Education*, *44*(1), 97–110. https://doi.org/10.1080/02602938.2018.1476670

Johnson, M., & Kuennen, E. (2006). Basic math skills and performance in an introductory statistics course. *Journal of Statistics Education*, *14*(2). https://doi.org/10.1080/10691898.2006.11910581

Learning Assistant Program: Faculty Resources. (2020). Retrieved from https://www.colorado.edu/program/learningassistant/faculty-la-resources

Liu, M., McKelroy, E., Corliss, S. B., & Carrigan, J. (2017). Investigating the effect of an adaptive learning intervention on students' learning. *Educational Technology Research and Development*, *65*(6), 1605–1625. https://doi.org/10.1007/s11423-017-9542-1

Lovett, J. N., & Lee, H. S. (2017). New Standards Require Teaching More Statistics: Are Preservice Secondary Mathematics Teachers Ready? *Journal of Teacher Education*, *68*(3), 299–311. https://doi.org/10.1177/0022487117697918

Minitab, L. (2019). What is a General linear model? Retrieved from https://support.minitab.com/en-us/minitab/18/help-and-how-to/modeling-statistics/anova/supporting-topics/basics/what-is-a-general-linear-model/

Mocko, M., Lesser, L. M., Wagler, A. E., & Francis, W. S. (2017). Assessing effectiveness of mnemonics for tertiary students in a hybrid introductory statistics course. *Journal of Statistics Education*, *25*(1), 2–11. https://doi.org/10.1080/10691898.2017.1294879

National Committee on Science, Engineering, and P. P. (2011). *Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline Committee on Science , Engineering , and Public Policy Policy and Global Affairs THE NATIONAL ACADEMIES PRESS*. https://doi.org/10.1007/s10273-011-1262-2

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). Common Core State Standards for Mathematics (Data and Statistics).

Otero, V., Pollock, S., & Finkelstein, N. (2010). A physics department's role in preparing physics teachers: The Colorado learning assistant model. *American Journal of Physics*, *78*(11), 1218–1224. https://doi.org/10.1119/1.3471291

Pande, P. S., Neuman, R. P., & Cavanagh, R. R. (2007). *The Six Sigma Way Team Fieldbook*. https://doi.org/10.1007/978-3-8349-9320-5_24

Primi, C., Donati, M. A., & Chiesi, F. (2016). A mediation model to explain the role of mathematics skills and probabilistic reasoning on statistics achievement. *Statistics*

Education Research Journal, *15*(2), 169–178.

Pyzdrowski, L. J., Sun, Y., Curtis, R., Miller, D., Winn, G., & Hensel, R. A. M. (2013). Readiness and Attitudes As Indicators for Success in College Calculus. *International Journal of Science and Mathematics Education*, *11*(3), 529–554. https://doi.org/10.1007/s10763-012-9352-1

Qlik. (2018). Data Literacy Survey Results. Retrieved October 4, 2018, from www.dataliteracy.info/application

Quality Council of Indiana. (2014). *CSSGB Primer* (3rd ed.). Terre Haute, IN: Quality Council of Indiana.

Reisener, C. D., Dufrene, B. A., Clark, C. R., Olmi, D. J., & Tingstrom, D. H. (2016). Selecting Effective Interventions to Increase Math Computation Fluency via Brief Experimental Analysis. *Psychology in the Schools*, *53*(1), 39–58. https://doi.org/10.1002/pits

Rumsey, D. J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education*, *10*(3). https://doi.org/10.1080/10691898.2002.11910678

Schunk, D. H. (2012). *Learning Theories: An Educational Perspective* (6th ed.). Pearson Education Inc. https://doi.org/10.1017/CBO9781107415324.004

Sharma, S. (2017). Definitions and models of statistical literacy: a literature review. *Open Review of Educational Research*, *4*(1), 118–133. https://doi.org/10.1080/23265507.2017.1354313

Skemp, R. R. (1978). Relational understanding and instrumental understanding. *Arithmetic Teacher*, *26*(3), 9–16.

VanDerHeyden, A. M. (2015). Approaches to RTI. Retrieved from http://www.rtinetwork.org/learn/what/approaches-to-rti

Wallman, K. K. (2006). Enhancing Statistical Literacy: Enriching Our Society. *Journal of the American Statistical Association*, *88*(421), 1. https://doi.org/10.2307/2290686

Weiss, D. M., & Belland, B. R. (2016). *Transforming Schools Using Project-Based Learning, Performance Assessment, and Common Core Standards*. *Interdisciplinary Journal of Problem-Based Learning* (Vol. 10). https://doi.org/10.7771/1541-5015.1663

Whitaker, D., Foti, S., & Jacobbe, T. (2015). The Levels of Conceptual Understanding in Statistics (LOCUS) Project: Results of the Pilot Study. *Numeracy*, *8*(2). https://doi.org/10.5038/1936-4660.8.2.3

Yolcu, A. (2014). Middle school students' statistical literacy: Role of grade level and gender. *Statistics Education Research Journal*, *13*(2), 118–131. Retrieved from http://link.galegroup.com.prxy4.ursus.maine.edu/apps/doc/A450999634/PROF?u=maine_orono&sid=PROF&xid=ec18d6ba

APPENDIX I: 2018 MAINE LEARNING ASSISTANT INTERVIEW TRANSCRIPT

About how many statistics boot camps did you conduct during the course?

I believe there were about 7, from around the second week of the course until the 6th week, twice a week.

About how long did each boot camp last for?

About 45 minutes from beginning to end, some folks heading out a little early.

What were some topics covered in the boot camps?

Central Tendency, Dispersion, Reading z- and t-charts, what a normal distribution acts like, and types of probability.

If you had to guess what fraction of the whole class attended the boot camps, what would be your estimate?

From 1/3rd of the class to a half, depending on the day. For a lot of folks the first week of the boot camps was definitely review, so those sessions on central tendency were probably the least attended

Did you notice if one major seemed to attend more frequently? (example: MET vs. bio/chem engineers…)

Was a pretty even cross section of the course, one thing I noticed though was that more METs seemed to come even if they weren't recommended by the Concept Inventory results. Several didn't even take the concept inventory, saying they were going to the boot camps regardless.

What parts of the boot camps do you think were especially helpful in encouraging student success in the course?

I think it really helped with the repetition and conceptual understanding that was very quickly covered in class or assumed already there. It focused in on these concepts and made sure students had a chance to go through them at a slower pace and increased depth, which is especially important in the case that they are encountering the concepts for the first time.

What parts do you think were not helpful, if any? Could you think of changes in boot camp design that could address this?

At first, the group worksheets at the end of the boot camp didn't seem like a good use of time, since everyone just kind of did them alone quietly. I felt like if that was going to be the case, that they may make better homework sheets than ones done within the camp. Once the students got comfortable with each other in the following weeks and were actually getting together collaboratively, that really improved their effectiveness, and I saw more why they were included.

While in the process of conducting the boot camps, did you make any changes regarding how the were conducted? If so, why do you think you made these changes?

When I started I wasn't really all that confident in my teaching, so I for the most part kind of read from the first worksheet as part of the lecture type part of the boot camp section. That didn't really resonate with me, I knew the students knew how to read! Over time it became more of a discussion than a lecture, about why the statistics concept is important, why engineers want to measure it, and maybe some benefits and limitations of using the statistics concept to choose a test or make a conclusion.

Did conducting the boot camps or being an MLA change any ideas you had about teaching or education?

I didn't know all that much about teaching, other than really small group tutoring I had done for 3 years prior. However, I did know from undergrad what types of teaching worked for me and which didn't. So when I started, I tried to focus on the ones I thought of as effective. I found out pretty quick that they were not universal, often you had to vary how teaching was presented in order to get more people on board and developing understanding

The MLA work really brought out a lot more of what I thought education could be like and how it can be made better. It even got me started with working with the RiSE Center, and helped me to decide to get more involved in engineering education.

BIOGRAPHY OF THE AUTHOR

Justin is a Master of Science in Teaching (MST) student, currently living in Rockland, ME. He graduated from the University of Maine’s Mechanical Engineering Technology bachelor’s program in 2018. Since enrolling at the university, Justin has also worked as a Research Technician intern in the research engineering department of Maine Maritime Academy, and at the composite manufacturer Compotech Inc. in Brewer, ME. During his undergraduate career, Justin tutored Technical Physics and Calculus courses for the University of Maine as a CRLA certified Experienced Tutor. He also provided homework help and study skills tutoring to students at Old Town High School and Leonard Middle School, free of charge through a nonprofit program he began at New Life Old Town church and youth center. He found out about the MST program after having the opportunity to work with the RISE center as a Maine Learning Assistant (MLA) in a trial run of the MLA program in a Lean Six Sigma course. Justin has recently begun teaching pre-engineering at Mid-Coast School of Technology in Rockland, ME. His research interests include applied statistics education, education methods in engineering, and career and technical education. He is a candidate in the Master of Science in Teaching degree in the Generalist option (Math and Physical Science) from the University of Maine in August 2020.