Deployment and Evaluation of an AI-Augmented Tutoring Application in a Classroom Setting

A Thesis
Presented to
The Academic Faculty

by

Joshua Taylor Hurt

In Partial Fulfillment
of the Requirements for the Degree
Master’s of Mechanical Engineering in the
The George W. Woodruff School of Mechanical Engineering

Georgia Institute of Technology
August 2021

COPYRIGHT © 2021 BY JOSHUA HURT
DEPLOYMENT AND EVALUATION OF AN AI-AUGMENTED TUTORING APPLICATION IN A CLASSROOM SETTING.

Approved by:

Dr. Julie Linsey, Advisor  
School of Mechanical Engineering  
*Georgia Institute of Technology*

Dr. Aldo Ferri  
School of Mechanical Engineering  
*Georgia Institute of Technology*

Dr. Wayne Whiteman  
School of Mechanical Engineering  
*Georgia Institute of Technology*

Date Approved: July 27, 2021
To education, to the future.

The hardest part of ending is starting again.
ACKNOWLEDGEMENTS

I’ll start by thanking the many professors that have assisted me in my journey through education, especially my advisor Dr. Linsey. Without her guidance, I would have never applied to the graduate program in the first place. I would also like to thank the many professors that I have worked with as part of the CHAMP grant, especially Dr. Hammond who has been instrumental in this whole process. Thanks to the members of my committee as well, who helped shape my understanding of engineering during my undergraduate curriculum.

I would like to thank my friends and family, whose emotional support gave me the strength to get through college. I would especially like to thank my parents, whose support got me through 19 years of education and enabled me to begin my graduate degree. Without them, I wouldn’t have had the bravery to further my education. I would also like to thank my wonderful girlfriend Yudi Liu, whose love helped me through some difficult times through the past few years.

Thanks to the many graduate and undergraduate students that I have worked with, especially my labmates Tim Sawchuck, Morgan Weaver, and Alex Murphy. Without their help and guidance my experience would have been entirely different. In addition, many thanks to the many students that participated in the study, your efforts have meant more than you could ever know. Finally I would like to thank the National Science Foundation for funding the CHAMP grant. As always, any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv  
LIST OF TABLES vii  
LIST OF FIGURES viii  
LIST OF SYMBOLS AND ABBREVIATIONS xi  
SUMMARY xii  

CHAPTER 1. INTRODUCTION 1  
1.1 Context and Reasoning 1  
1.2 Research Questions and Objectives 4  
1.3 Structure of the Thesis 7  

CHAPTER 2. BACKGROUND 8  
2.1 Mechanix 8  
2.2 Importance of Free-body Diagrams in Engineering 11  
2.3 Existing Online Homework Systems 12  
2.4 Feedback in Engineering Education 14  
2.5 Creativity and Hands-On Participation in Engineering 15  
2.6 Concept Inventories 17  

CHAPTER 3. METHODOLOGY 19  
3.1 Experiment overview 19  
3.2 Instrument Distribution 23  
3.3 Truss Design Problem Distribution 26  
3.4 Control Participant Submission Coding 32  
3.5 Pilot Study: Dynamics Distribution 34  

CHAPTER 4. RESULTS 36  
4.1 Mechanix Homework Distribution Results 36  
4.1.1 FBD Problem Set Scores 37  
4.1.2 Mechanix Assignment Truss Problem Set 39  
4.2 Exam Results 41  
4.2.1 University 1 Exam Results 41  
4.2.2 University 2 Exam Results 43  
4.2.3 University 5 Exam Results 44  
4.3 Concept Inventories 46  
4.3.1 Force Concepts Inventory 46  
4.3.2 Statics Concepts Inventory 47  
4.4 Pilot Study: Dynamics Problem 51  
4.5 Truss Design Problem Results 56  
4.5.1 Truss Design Problem Pilot Study 56  
4.5.2 Expanded Problem Distribution 62
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>The overall distribution of usable data at the universities involved in the study. Participation at each university varied from semester to semester, with lost data sets due to COVID-19.</td>
<td>21</td>
</tr>
<tr>
<td>Table 2</td>
<td>The following table shows the assignment of students and Sections to the different experimental conditions, as well as notations of the different factors that affected data collection during the study.</td>
<td>22</td>
</tr>
<tr>
<td>Table 3</td>
<td>An overview of the traditional homework systems used at the participating University</td>
<td>24</td>
</tr>
<tr>
<td>Table 4</td>
<td>Initial deployment of the Truss Design Problem was isolated to University 1, with eventual full deployment beginning during the Fall 2020 semester. Two problem types were distributed, the Low Load and the High Load truss problems.</td>
<td>27</td>
</tr>
<tr>
<td>Table 5</td>
<td>The following table details the types of feedback given to participants assigned to the Control conditions. Some of these types of feedback are impossible to receive through the Mechanix application</td>
<td>33</td>
</tr>
<tr>
<td>Table 6</td>
<td>The following table details the types of feedback given to participants assigned to the Control conditions. Some of these types of feedback are impossible to receive through the Mechanix application</td>
<td>68</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1  The User interface prompts the user to draw the FBD and all associated parts, automatically updating the entry fields as the user draws the diagram. 10

Figure 2  As with the general bodies problems, Mechanix prompts the user to draw the FBD of the system. The UI automatically provides entry fields for the internal force members of the method of joints problem, providing feedback on any errors. 10

Figure 3  One of the variants used in the study, which was deployed to the participating Aerospace courses taught by the University 2 instructors. 28

Figure 4  The different variations in the distributed problem statements can be seen above. Not pictured is the variation used at Univ. 3, which altered the problem to use the U.S. customary units. 30

Figure 5  Example of a drawn solution. The truss design problem asks students to draw the proposed truss, followed by the input of the proposed geometry and then finishes with the application of the loads. 31

Figure 6  One of the Mechanix problems developed which uses Dynamics concepts, which was deployed to Dynamics classes at University 3. 35

Figure 7  The data summarized above is taken from Fall 2019, Spring 2020, and Fall 2020. All schools in the participating data sets ran the general bodies assignment as the first Mechanix activity, with no statistical difference between the Control or Mechanix conditions. 38

Figure 8  Comparison of the data collected at the universities during Fall 2019, Spring 2020, and fall 2020. Results suggest continued usage of the application assist student performance using Mechanix. 39

Figure 9  Results of the Mechanix intervention at University 5 suggests that the usage of the Mechanix application has a learning curve. 40

Figure 10  On all three exams, both the Mechanix and Control conditions performed equally. A trend occurred following the Mechanix intervention (Exam 2), suggesting the Mechanix application may have impacted student understanding. 42
Figure 11  Statistical difference between the conditions on the experimentally controlled question. While students did not show a statistically higher grade on the exam overall, this result suggests that the application improve the Mechanix condition scores.

Figure 12  The overall exam scores at University 2 (collected from Fall 2018, Fall 2019, and Fall 2020) reflect the other instruments distributed at University 2, with no indication that the Mechanix application harmed student understanding of the material.

Figure 13  Through the analysis of the exam scores at University 5, it was discovered that the time of day had a larger impact on student performance than the introduction of the Mechanix application.

Figure 14  FCI results across each semester show no statistically significant difference between the Mechanix and Control conditions.

Figure 15  Like with the FCI, the pre- and post-semester distribution of the SCI showed no significant difference between the Mechanix and Control homework conditions at all participating schools.

Figure 16  These results show that certain questions are more commonly answered correctly than others, possibly suggesting that certain questions are easier.

Figure 17  Following the reorder of the SCI, all had marginal improvements, with the results continuing to reflect previous results, except at University 5.

Figure 18  Through comparing the original SCI results and the reorder SCI, it was discovered that both the Control and Mechanix participants performed better on the reorder survey.

Figure 19  The results from the regrading effort and resulting analysis indicates that the administered exams during the Fall 2018 and Fall 2019 semesters may not be indicative of the learning outcomes of the Mechanix participants.

Figure 20  The Mechanix intervention had no impact on the results on the aDCI. This is consistent with Concept Inventory scores at other universities, although the conditions during data collection are expected to have affected the results.

Figure 21  The figure above shows the trivial solution of the low-load design problem.
Figure 22  Results of the Pilot distribution of the Truss Design problem indicated that the problem statement did not present enough of a challenge.

Figure 23  Overall participation during the shift to online instruction caused a general lack in participation in the Truss Design Problem.

Figure 24  A comparison of the total number of the mean total answers submitted per user on the Truss Design mode, showing that the Experimental condition submitted a statistically different number of submissions than the Control condition.

Figure 25  A comparison between the mean calculated load the submitted trusses could handle, showing that the spring 2020 Variation of the problem was solvable and that the experimental condition trended towards having better results.

Figure 26  The results at University 1 seem to suggest that in high-stress situations, the Mechanix application does not harm student performance.

Figure 27  The mean calculated load shows that the difference in the experimental conditions is in favor of the Mechanix condition.

Figure 28  The breakdown of the errors made by students by condition shows that the Mechanix application reduced both conceptual errors and the overall rate of mistakes made through the Mechanix application.

Figure 29  Shown in one figure, both the University 4 and University 5 experimental conditions are consistent with previous submission counts seen in the pilot distribution at University 1.

Figure 30  By Mirroring the low load trivial solution about the X-axis, a solution can be achieved that exceeds the high load problem statement. This new trivial solution was not submitted at the same rate as the low-load trivial solution.

Figure 31  first-time users of Mechanix tends to indicat a learning curve associated with the Mechanix homework system.
LIST OF SYMBOLS AND ABBREVIATIONS

FBD    Free Body Diagram
SCI    Statics Concepts Inventory
FCI    Force Concepts Inventory
aDCI   abridged Dynamics Concepts Inventory
CI     Concepts Inventory
SUMMARY

Many advocate for more opportunities throughout the curriculum for engineering students to practice their design and creativity skills. However, the large amount of time required for grading open-ended design problems as well as provide formative feedback on homework problems generally prohibits the extensive use of design problems. This restricts the type of assignments that students often receive, as well as the level and quality of feedback provided on homework problem sets. This thesis presents a study of the impact of Sketch Mechanix within introductory mechanics classes at multiple universities across the United States. The application, Sketch Mechanix, is an AI-augmented sketch recognition homework tool used to distribute traditional homework problem sets and open-ended design problems.

This application has been developed to reinforce the conceptualization, development, and solution of engineering systems through idealized free-body diagrams. Mechanix provides the user with a drawing space and uses sketch recognition algorithms to automatically detect general body FBDs and planar 2Dtrusses, providing tutoring towards correcting mistakes that may have been made. The study compared the academic impacts of Sketch Mechanix by dividing student participants at the multiple universities into a Control (online or pen-and-paper) condition or Mechanix condition, and asked them to complete a series of homework problems. The Control condition uses only the traditional homework system used in that class section, while the Mechanix condition completes a select set of homework problems using the Mechanix application. Participants were also
asked to complete an open-ended design problem in addition to the standard homework problems.

Conducted from Fall 2019 to Spring 2021, over 500 student participants have completed the problem sets, with 300 completing the design problem. Results suggest that the scaffolded tutoring built into the Mechanix application supports student learning, increasing student learning in the short-term. This can be seen through the short-term homework and test achievement of the Mechanix condition outpacing their Control condition peers. Additional results from the truss design study suggest that using the Mechanix application during the solution process enabled participants to explore more of the solution space. Specifically on the design problem, Mechanix participants produced truss solutions that exceeded the specific design criteria, while also producing final solutions that had an overwhelmingly lower rate of errors than their peers in the Control condition.
CHAPTER 1. INTRODUCTION

1.1 Context and Reasoning

It is common for large entry-level courses to use online homework systems, which provides little feedback on incorrect answers to manage the high volume of assignments and grading, or paper-based homework with limited and delayed feedback. This trend is especially pronounced in today’s early engineering education. Online homework systems have become the norm for many students and teachers throughout the United States. The introduction of online homework systems and a shift towards virtual education have caused the traditional method of hand-worked paper submissions harder for modern engineering classes. Coupled with the lengthy requirements to grade paper submissions, the level of feedback and mediation from instructors in engineering education has become more difficult, especially when regrade efforts may be a concern [1].

While the shift to online homework does not have significant adverse side effects on student understanding of the overarching concepts, it has been shown that the limited and feedback coupled with poor instruction can lead to misunderstandings in methods employed to solve the problems, especially among college students [2]. One essential concept whose value is diminished due to this shift in education and large class sizes in early engineering courses is the creation and usage of free-body diagrams (FBDs). The instant feedback provided by online homework systems is impactful [3], but they do not actively engage a student’s ability to accurately create FBDs in all situations. This decrease in FBD learning outcomes has not been explored in past research to the researcher's best knowledge.
Evidence suggests that the usage of FBDs is correlated to an increased understanding of Mechanics concepts, with a strong connection between usage of the correct implementation of FBDs and the solution of statics and dynamics problems [4]. The concern is that online homework applications do not fully reinforce the creation of FBDs. While students are capable of learning complex concepts in large classes and remote settings using traditional online homework methods, it is challenging to provide open-ended problems to students in these settings. The lack of support for these problems in remote learning has created a need for an online application that provides students with a platform to submit both their solutions and the associated FBDs easily through instant feedback provided to the students and automatic grading.

Homework distribution methods are not the only factor in student understanding of the material. The type of feedback has been shown to have an impact on student achievement, with formative and immediate feedback being the most constructive form of feedback [5]. This type of feedback is often missing in traditional homework systems and large classes, or greatly reduced based on the kind of homework system used in the classes. Because of these gaps in current instructional methodology, there have been calls by engineering educators to improve upon the existing deficiencies in education [6]. Major challenges exist related to the implementation of improved homework systems, both due to the existing time delay difficulties related to feedback in paper-and-pencil homework and the lack of constructive feedback in impartial online homework systems.

Furthermore, the ability for instructors to provide students meaningful opportunities to express scientific creativity in today’s introductory engineering courses is challenging, as these large class sizes greatly diminish an instructor's ability to provide
meaningful feedback on traditional creativity-based assignments. The inclusion of creativity in teaching curriculums is an essential part of students' education, with creativity being described as a critical part of an engineer’s training [7]. Creative implementation of ideas is known to have the greatest impact on long-term retention of knowledge, with the highest level of Bloom's taxonomy being creativity and creating. This suggests that students have the deepest understanding when they can apply a concept creatively [2].

Despite the importance of creativity in education, it is often difficult to provide students with meaningful ways to express their creative skills. Large class sizes in formative introductory courses and significant changes in educational strategies have made the traditional method of project-based creativity assignments increasingly challenging to distribute and provide meaningful feedback. This shift in education makes the traditional hand-worked paper submissions more challenging for modern engineering courses, forcing a change to wide usage of online homework methods.

One such method for introducing creative problem solving into engineering education and reinforcing graphical representation settings is distributing open-ended design problems in early engineering classes. These types of problems enable participants to apply concepts that they have learned in a creative capacity. While these problems can provide students with a creative application for the material they have learned, the grading and feedback process associated with these problems can limit an instructor's ability to evaluate open-ended design assignments. In support of open-ended design problems, it has been shown that active thinking and application of knowledge — as opposed to passive listening — creates an environment where students can more actively learn about the complex concepts presented during a class [8].
As a possible solution to these challenges in engineering education, an application has been developed to provide a source for instant formative feedback. Sketch Mechanix is an online AI-augmented FBD tutoring application, which has been created to provide personalized feedback. The application provides users with a space to present an FBD related to the problem, while sketch recognition algorithms provide real-time feedback on the student-drawn submissions. The intention is to build an improved scaffolded learning experience for early engineering students. The application supports general FBD systems, planar truss systems, as well as architecture for open-ended truss design problems.

This thesis presents the educational impacts of Mechanix, a platform designed to provide immediate, as well as constructive, and personalized through an online platform. This thesis focuses on the deployment of the application and the evaluation of the educational merits of the Sketch Mechanix at multiple universities across the United States. This deployment has involved a tiered rollout of the problems created in the Mechanix application, using a combination of pre-existing problems and problems by members of universities participating in the study. The creation of the software has been handled by the Sketch Recognition Lab at Texas A&M University.

1.2 Research Questions and Objectives

This thesis explores the impacts of using the Mechanix platform to distribute both standard problem sets and open-ended design questions in Statics and Dynamics courses. These problem sets focus on developing a student’s ability to correctly draw free body diagrams, with the focus being the overall difference in student achievement between participants who use the Mechanix application and those that use traditional homework
methods such as online systems or paper homework. The research questions acted as the focus of the study, and herein this thesis. The final questions focus on the impact of the distributed instruments on student understanding and achievement. The thesis investigates the following questions:

*RQ1*: To what extent does the usage of an AI-augmented tutoring system (Sketch Mechanix) improve student learning outcomes in Statics and Dynamics courses?

*RQ2*: How does the use of Mechanix affect the conceptual knowledge gains in a Statics course, as measured using established concept inventories?

*RQ3*: What are the impacts of using an intelligent tutoring software to solve an open-ended creative design problem?

These questions aim to investigate the overall academic impact of the Mechanix application, specifically if using the application improves learning outcomes of students. From past work using previous versions of the application, it has been shown that Sketch Mechanix has an impact on the educational outcomes of high achieving students [9-11]. The application also provides additional motivation to students that are referred to as at-risk of leaving engineering education [12, 13]. The driving hypotheses of this study focuses on the impact of Mechanix on student understanding. Past research has shown that the Mechanix application has a positive influence on student understanding of the material. Specifically, it is expected that the usage of the Mechanix application, and the immediate feedback provided by it, will allow participants to have a greater understanding of concepts that they may have misunderstandings of. This manuscript focuses on the academic effects of the current application.
**RQ1:** The expectation is that students who complete assignments using the Mechanix application for the related homework assignments will show improved understanding of the assigned material compared to their peers who complete similar homework problems using traditional homework methods. This is due to the built-in scaffolded, immediate, and formative feedback and evaluation of the FBDs drawn in the Mechanix application. This hypothesis is expected to be reinforced by both the overall homework grades and the results from the collected exam questions from each university.

**RQ2:** The hypothesis is that students who use the Mechanix application will show equal performance on the peer-reviewed concept inventories (CI) distributed as part of the study. This is hypothesized due to the difficult nature of these concept inventories, showing that any minor differences in academic performance between the students who use the Mechanix application and their peers who use the traditional method will only have minor impacts on the highly conceptual nature of the CI’s. The second expectation is that participants from the Universities involved in the study will also show comparable performance, regardless of the school that the instrument is distributed.

**RQ3:** The expectation is that using the Mechanix application for the open-ended design problem will enable participants to explore more of the design space. The instant feedback, the distribution method, and the time-saving aspects of the application are expected to enable the participants to solve the problem and provide immediate reinforcement on skills that may be lacking during the solution process.
1.3 Structure of the Thesis

The following manuscript is ordered as follows. The manuscript begins with a description of the study’s reasoning (Background), followed by the methodology used in the deployment of the Mechanix application and the associated instruments used to evaluate the application’s academic impacts (Methods). From there, the results of the instrument distribution are described (Results), with a corresponding discussion of the impacts of the analyzed instruments used in the study (Discussion). The final statements on the deployment and impact of the Mechanix application, as well as future steps that are, conclude the paper (conclusion).

The manuscript is ordered to explain the steps taken in deploying and evaluating the impact of the Mechanix application on a diverse set of student participants. Each chapter is organized to define and describe the reasoning, deployment, and evaluation of the instruments used as part of the study. The outcomes are divided into two sections: the General Impacts of Sketch Mechanix and the results of the Creative Truss Design problem distribution. The general Impacts of Mechanix follows the distribution of the standard Mechanix problems and the associated results from the collected exam questions and the peer-reviewed Concept Inventories (Cis). The remainder of the manuscript focuses on the deployment and analysis of the Creative Truss Design problem and the impact that this design problem has on the achievement of the student participants. This problem, and the associated reasoning, are presented as an example of a modern solution to the lack of creativity-based assignments in theory-based engineering curriculums.
CHAPTER 2.  BACKGROUND

New engineers often have inadequate experience sketching by hand for the communication of ideas [14], which has been shown to create a shallow understanding of how to communicate complex engineering concepts, thereby decreasing the innovation capabilities of engineers in industry [15]. Modern educators see a need to strengthen these communication and sketching skills [6, 16]. In today’s engineering education curriculum, hand drafting has been replaced with education in computer-aided design (CAD). Some indications show that while this does not hinder student understanding of concepts [17], it muddles the true impact of CAD and reduces the emphasis on hand-sketched ideas and communication [18]. Sketching has also been “understood to act as a key factor for creative expression” and “one of the most effective visual thinking.” [19]. FBDs are a form of sketching in early engineering classes and an important aspect of a students’ education. Furthermore, it can be argued major shifts in education also affect student understanding of sketching and the usage of FBDs, due to the shifts in educational assessment related to the requirements of online homework systems.

2.1 Mechanix

Mechanix has been envisioned and developed as an online platform where students can learn complicated engineering concepts by solving problems on a sketch-recognition interface. This platform enables students to sketch their free body diagrams on a virtual interface, and the system recognizes the sketches using an AI algorithm. While students solve problems on the virtual interface, Mechanix provides them immediate formative feedback, which the literature shows to be an effective learning tool [20]. The platform
also uses scaffolded learning, an effective way to understand complex concepts [21]. The application's architecture also provides instructors with the opportunity to easily deploy an open-ended design problem, with the overhead of a standard homework problem due to the online distribution platform and the auto-grading features built into the application.

Through Mechanix, students complete homework problems using the AI-augmented system, in which they create drawn FBDs as if they were drawing them for a hand-written solution(Figure 1 and Figure 2). As they create parts of the FBD, such as adding new truss members or increasing the number of loads, the UI automatically populates with the next step in FBD creation and adaptively provides a scaffolded solution approach through the updated inclusion of answer boxes for the student to fill with their proposed solutions. Past evaluation of the application shows that it is as impactful as traditional homework methods (paper and pencil) while assisting students considered at-risk for leaving engineering education in building on their ability to draw FBDs [10-13].

The application has built-in sketch recognition support for general body FBD problems and method of joints truss problems, as well as support for vector mechanics problems. Mechanix asks a participant to draw the FBD associated with a given problem, with problem order being presented to produce a Scaffolded learning experience in introductory Statics and Dynamics classes. The application provides feedback related to both errors in the drawn FBDs, as well as the proposed solutions to the problem, allowing more formative feedback on mistakes a student may make during the solution process. The application recently has been expanded to include an open-ended design problem[22], with increased recognition for more components commonly see in FBDs.
Figure 1: The User interface prompts the user to draw the FBD and all associated parts, automatically updating the entry fields as the user draws the diagram.

Figure 2: As with the general bodies problems, Mechanix prompts the user to draw the FBD of the system. The UI automatically provides entry fields for the internal force members of the method of joints problem, providing feedback on any errors.
2.2 Importance of Free-body Diagrams in Engineering

Through comprehensive studies of student exams in different types of physics-related courses, introductory statics courses, and dynamics courses, it has been shown that the accurate sketching and usage of FBDs are impactful on an understanding of physics-related concepts at all levels of education [23]. In support of this, many statics textbooks introduce the concept of FBD sketching in the opening chapters [24-26]. The concept is of such importance that more conceptually difficult textbooks dedicate large portions of their opening chapters reviewing the creation methods and usages of FBDs, such as the industry-standard *Shigley’s Mechanical Engineering Design* [27] as well as *Mechanics of Materials* by Gere and Goodno [28]. While students may not know the immediate positive impact of FBDs on problem conceptualization, this critical tool has been shown to have a significant role in engineer’s ability to understand difficult material [29]. It has also been shown that the usage of FBDs for solving physics-related problems correlates with a better understanding of the problems that are asked [30]. Despite this, new engineers often have inadequate experience preparing sketches by hand to communicate ideas [14, 31] which is especially true for FBDs.

While it can be shown that students can achieve a complete understanding of difficult concepts while using online homework, the form of feedback provided to instructors from these websites may not inform them of deficiencies their students may be experiencing [32]. In cases where the problems in the online homework system are poorly created, students are typically encouraged to employ a “plug-and-chug” method of brute force solving instead of internalizing their errors and correcting them [32]. Furthermore, it has been shown that a large variety of homework problems produce conceptually rich
conversations between members of a class section, both during classroom discussions and outside of classroom settings [33]. These conversations can be marred by implementing these online homework systems, especially when poorly formatted problems create negative attitudes within the students [34].

2.3 Existing Online Homework Systems

Homework applications such as WileyPlus have been shown to be useful in higher levels of engineering education, with the caveat that freshmen college students often oppose the usage of these impartial systems due to inexperience with these types of systems [35]. Furthermore, online homework systems, such as Wileyplus and Mastering Engineering, only passively support the practice of creating FBD for the solution of the problems by allowing participants to submit their work, with no feedback unless instructors manually review it. It has long been questioned whether or not these online applications effectively teach engineering students to clearly present their work, especially related to the creation of FBDs [36]. It has been shown that the usage of these online homework systems has advantages and disadvantages. Analysis of the impact of these homework systems on a wide variety of students shows that the loss of motivation, oversimplified answer structure, and railroading present in decision-making built into the usage of these applications, acts as a major drawback related to student engagement [37]. One major advantage comes in the instant provided by the online problem homework as well as the mass distribution of problems built into online homework systems.

Some multi-disciplinary studies on the impact of these online homework systems have shown that while online homework can be correlated with good concept practice, it
also cannot be demonstrated that this practice created through online homework are better than traditional pen-and-paper homework [38]. Within the literature, there is general support that online homework systems have a place in college education [39, 40], with students being more willing to pay for online homework systems that provide more educational support [41]. Despite the support for the overall effectiveness of online homework systems, there is a lack of support regarding any losses that the usage of these applications may have on student understanding and usage of FBDs.

Some applications have been created to provide students with the opportunity to visualize and solve graphic representations of engineering problems and further explore expanded engagement in physics concepts, especially related to the usage of FBDs. Most of these applications come in the form of drag and drop palate-based systems, which employees a set number of pre-programmed objects and asks a participant to place the pre-generated objects in the correct locations. While this solution method may increase student engagement, it reduces the cognitive load on a student and may restrict a student’s ability to move a concept from working memory to long-term memory [42], as they simplify the process to forms of fill in the blank style of questions. Some of these systems that have been created by different educational providers include WinTruss [43], McGraw Hill Connect Engineering [44], Bridge Architect [45], and Andes physics tutoring system [46]. These homework systems represent an advancement in instruction. However, the simple problem types and the palette-based actions reduce the need for a student to shift information they are learning from working to long-term memory or retrieve aspects from long-term memory and apply it [47].
Other applications have been produced by educational researchers. One of these applications is known as Newton’s pen, which allows users to draw the FBD. However, the application does so by constraining users to do so create the FBDs in a specific order [48, 49]. Another system created by educational researchers at Arizona State University focuses on the development of Circuits Tutoring through an interactive system that is reminiscent of a palate-based U.I [50]. Another palette-based educational tool that has been created is the TrussMe! Phone app [51], which experiences the same downsides as other palette-based solution methods. Yet another application that is the work of Roselli et al. at Vanderbilt University has resulted in a 2-D FBD assembly tool, although the application follows an ordered scheme similar to Newton’s pen. This reduces the amount of input needed by the participant by restricting the need for the user to conceptualize the overall process [52, 53]. None of these applications provide a natural interface for hand sketching or open-ended design problems, which is something that Mechanix provides.

2.4 Feedback in Engineering Education

One of the benefits of online homework comes in the form of immediate feedback, which plays a crucial role in improving student understanding [54]. Most traditional pen-and-paper homework methods provide crucial amounts of feedback; however, this feedback is delayed by a significant period of time and may be minimized in large classes. This gap in available feedback creates a period where students may be unaware of their mistakes, creating problems for future understanding of the material. A meta-analysis concluded that immediate feedback provides greater benefits than the delayed feedback that often occurs [3], with the results being especially consistent within undergraduate populations [55]. It has also been shown that instant feedback has a positive effect on
student anxiety, significantly reducing the overall anxiety students feel related to the subject [5]. Feedback comes in multiple forms, with two of the major types being formative and summative feedback. Formative feedback is defined as feedback that attempts to alter a learner’s thinking or behavior during an assignment [56], while summative feedback is described as feedback provided at the end of an assignment or course [57]. Comparisons of the two types of feedback have shown that formative feedback is more impactful than summative feedback, especially related to the completion of homework assignments [58].

In addition to the type of feedback, a method known as scaffolded learning has been shown to increase student understanding by breaking down a problem into more straightforward steps and providing feedback at each step [59]. This scaffolded feedback has been shown to support student understanding of difficult concepts, essentially activating and enabling a students’ ability to create solutions related to a given problem statement [21]. This type of feedback is especially effective when students are encouraged to present more than one solution to the problem, which has been shown to increases homework achievement [60], supporting multiple attempts on online homework. Furthermore, automated feedback and scaffolded learning have been shown to reduce the overall load on already time-taxed instructors [61], supporting the need to provide students with scaffolded opportunities of a formative nature.

2.5 Creativity and Hands-On Participation in Engineering

Evidence suggests that students understanding of complex concepts is deepened when they are provided with opportunities to creatively employ these concepts in a meaningful way [62]. However, the increase in student enrollment in higher education has
impacted educational strategies and outcomes [63], with a cursory glance at engineering curriculums reflecting this loss in the creative application of knowledge. The creative application of knowledge, and the associated abstraction related to these methods of assessment, are known to be a solid indicator of student understanding of these complex objects and have been shown to increase learning effectiveness [64].

It has been shown that hands-on opportunities and open-ended design problems enable students to better understand complex topics [65]. These types of tasks have been shown to be so influential that some curriculums have been reformatted to encourage more of these types of tasks [66]. Further, educators believe that the pre-existing deployment of laboratory settings has lost its educational focus and meaning [67], which suggests that while students may get some hands-on reinforcement, it may be viewed as a busy class with no educational merit. While the merit of some classes, such as Capstone-style courses, are known, they often occur towards the end of an engineering student’s undergraduate degree, which fails to reinforce student understanding during their entire education. While Capstone and lab courses have been shown to teach hands-on skills effectively [65], integrating hands-on tasks into multiple courses could further students’ understanding of complex concepts [68].

Hands-on design activities have been shown to be difficult to integrate into standard classes, as the time input required to correctly deploy them often forces them to be contained in lengthy standalone class course sections [69], as part of a major assignment, or forces courses to be a transition to Flipped Classrooms [70]. Time constraints in instruction often necessitate introducing these problems as part of a lengthy homework or project assignment. One of the most common assignments is the broad deployment of the
“Mouse-Trap Car” in K-12 and early undergraduate courses, which prompts students to use only the mechanical advantage of a mouse trap to create a car that can travel the farthest distance [71]. Another common design activity in early engineering education is model rocket competitions, which often take a considerable amount of time or an entire course dedicated to it to implement in a meaningful way [72, 73]. Assignments like this have been shown to create a lasting impact on student understanding at all stages of education, although the time required to deploy these assignments necessitates a lengthy grading and evaluation process. The deployment and grading process further restricts larger deployments of hands-on activities, which might be overcome by integrating project-based problems into online activities.

2.6 Concept Inventories

It should be mentioned that concept inventories (CI) have been shown to be valuable tools in evaluating students' understanding of concepts and, as such, are used as evaluation tools within the testing of the Mechanix application. One of the most well-known concept inventories, the Force Concepts Inventory (FCI) [74], has been shown to be an effective indicator of general Newtonian physics knowledge [75]. While this concept inventory is a well-known indicator of physics knowledge, a need for more concepts inventories that target specific conceptual knowledge has spurred the creation of focused concept inventories [76]. One of the newer concept inventories that have been developed is the Statics Concepts Inventory (SCI) [77]. The FCI and SCI represent peer-reviewed methods of ensuring participant understanding of essential engineering concepts. Another CI that has been shown to be representative of early engineering students' knowledge is the Dynamics Concepts Inventory [78], which has been subsequently modified and
streamlined for faster implementation. This abridged Dynamics Concepts Inventory (aDCI) is shown to be “more amenable to in-class implementation than the full DCI,” as it reduces the overall time it takes for students to complete the full inventory [79].
CHAPTER 3. METHODOLOGY

3.1 Experiment overview

The evaluation of the educational merits of Mechanix has been conducted through the widespread deployment at multiple universities across the United States. The schools were selected based on the desire to expand participant diversity, as previous studies on Mechanix focused on highly selective schools. Prior studies showed preliminary educational merits of Mechanix, especially amongst struggling students [10-13]. Mechanix was deployed at the four large public universities (University 1, University 3, University 4, and University 5) and one teaching-focused private university (University 2). These universities were selected based on the student body present at the university, as a more diverse subset of student participants using Mechanix would improve results of the study on the academic impacts of Mechanix by capturing more participants' responses to the academic merits of the application. Previous studies on Mechanix have shown that the application affects students who are at risk of leaving engineering who, due to inadequate preparation by their high schools, failed the course during a prior semester. The breakdown of the overall Mechanix problem distribution can be seen in Table 1 below, noting that some data sets were collected by not analyzed.

The distribution of problems followed the creation of new Mechanix problem sets, with more problems being introduced to students as the capabilities of Mechanix were improved. The homework problem sets were named based directly on what the focus of the problem set was, with the FBD homework focusing on the interpretation of a general body similar to a physics situation, and the Truss Homework focusing on solving...
truss systems using the method of joints. The final problem type that was distributed to participants was the truss design problem, which was created as a form of an open-ended design problem specifically for the creative application of difficult concepts. Table 2 shows the breakdown of the experimental conditions assigned to the participating class sections in more detail, as well as confounding issues that occurred during the semesters of distribution.

This thesis compares the educational merits of the application by contrasting student achievement between participants assigned to the Control or Mechanix conditions. The usage of Sketch Mechanix was integrated directly into the participating class sections at all five universities, with only select homework assignments being assigned through the Mechanix application. Participants assigned to the Control condition were asked to complete a series of CI’s, in addition to completing homework assignments through the normal homework method (paper or online systems) employed by that class section. Participants assigned to the Mechanix condition completed most homework assignments using the traditional homework method but were asked to complete a few problem sets using the Mechanix application, which used the same problem set as the traditional homework distribution. Homework problems from participants assigned to both conditions were collected for analysis. On all homework assignments given to both conditions, participants were given three attempts on each question to reduce the fear of failure.
Table 1: The overall distribution of usable data at the universities involved in the study. Participation at each university varied from semester to semester, with lost data sets due to COVID-19.
Table 2: The following table shows the assignment of students and Sections to the different experimental conditions, as well as notations of the different factors that affected data collection during the study.

<table>
<thead>
<tr>
<th>University 1</th>
<th>F 2018</th>
<th>S 2019</th>
<th>F 2019*</th>
<th>S 2020*</th>
<th>F 2020*</th>
<th>S 2021*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assigned as</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>University 2</th>
<th>F 2018</th>
<th>S 2019</th>
<th>F 2019*</th>
<th>S 2020*</th>
<th>F 2020*</th>
<th>S 2021*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assigned as</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>University 3</th>
<th>F 2018</th>
<th>S 2019</th>
<th>F 2019*</th>
<th>S 2020*</th>
<th>F 2020*</th>
<th>S 2021*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assigned as</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>University 4</th>
<th>F 2018</th>
<th>S 2019</th>
<th>F 2019*</th>
<th>S 2020*</th>
<th>F 2020*</th>
<th>S 2021*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assigned as</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mechanix</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>University 5</th>
<th>F 2018</th>
<th>S 2019</th>
<th>F 2019*</th>
<th>S 2020*</th>
<th>F 2020*</th>
<th>S 2021*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assigned as</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mechanix</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* COVID-19 affected data collection
1 Wild Fires Interrupted data collection
2 Winter Storms Interrupted data collection
3 Instructor Change During the Semester
Students were randomly assigned to either the Control or Mechanix conditions. Universities with smaller class sections were randomly assigned to a condition at the section level rather than assignments given to individuals. In contrast, sections with larger populations divided the participating population within a class into a “split” Mechanix and Control conditions. The overall intention was to maximize the number of “split” sections, as slight deviations in time of day, instruction methods’ or curriculum between semesters may have an unintended and unavoidable impact on the results. All participants received instruction in the same manner.

3.2 Instrument Distribution

As part of the study, students were asked to complete a series of pre- and post-semester diagnostics, including the Statics Concepts Inventory (SCI), an abridged Force Concepts Inventory (FCI), and the abridged Dynamics Concepts Inventory (aDCI) (See Table 2). The assigned instruments were used to compare a student’s understanding of the material at the beginning and the end of the class, comparing the effects of the Mechanix application on student understanding of essential engineering concepts. The merits of these CI’s were discussed at length in the background section of the manuscript, with the deployment of the aDCI as part of the discussion of the dynamics pilot study.

Each CI was implemented with a specific goal in mind. The FCI was implemented with the intention of using it as a covariant in the analysis of the more targeted CI’s. This instrument was deployed with the expectation that participants would have a solid conceptual grasp on the Newtonian physics covered by the instrument, and in an attempt to reduce the load of instruments distributed to participants, this survey was reduced in
length. The question count was reduced from the standard 30 questions to 8, with questions specifically selected to cover a subset of important concepts directly related to static analysis. Distribution of the SCI was expected to yield more pertinent results, as the CI targeted concepts directly taught by the classes. For this reason, the full 27 questions were deployed as both homework and classroom activities.

In addition to the pre- and post- semester diagnostics, homework results from the specific problem sets were collected from both the Control and Mechanix groups. Identical homework problems were distributed using the traditional homework system and Mechanix, with the results being an immediate comparison on the impact of the Mechanix system. The type of homework problem varied by school, with each university using the basic system already in use at that university; these systems can be seen in Table 3. Incidentally, each university used a different homework system across the duration of the experiment: For the Mechanix distribution, students at University 1, University 2, and University 4 completed an initial general bodies FBD assignment and a method of joints truss analysis assignment as part of the study, while students at University 5 completed only the method of joints truss assignment. The deployment of the Mechanix application at University 3 is detailed in the “Pilot Dynamics Distribution” section below.

Table 3, An overview of the traditional homework systems used at the participating University

<table>
<thead>
<tr>
<th>Homework Type</th>
<th>University 1</th>
<th>University 2</th>
<th>University 4</th>
<th>University 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WileyPlus</td>
<td>MasteringEngineering</td>
<td>Paper and Pencil</td>
<td>Canvas Quizzes</td>
</tr>
</tbody>
</table>

24
The effects of the Mechanix application were captured through the repeated distribution of homework assignments through the application, conducted through the initial deployment of a general bodies FBD assignment and a subsequent method of joints truss analysis assignment. The exact number and order of the problem set deployed varied by school, with the distribution of problems varying based on the internal curriculum of the classes of each school. The overall distribution of problems can be seen above in Table 1. The timing of the classes occurs directly based on the curriculum of the class, and therefore was often distributed over the course of a few weeks during the lecturing of the associated topic. Student results from both experimental conditions were removed from the analysis if they showed low performance, as it was assumed that these students did not make attempts at completing the problem sets. Participants were given a short demonstration of the usage of Mechanix prior to the assignment. Participants were also provided a series of tutorial videos for further explanation of how to use Mechanix and a short instruction handout on the application, both of which are contained in Appendix A.

To supplement the homework results, select exam questions were collected from participants of the study. Student achievement on these exams is used to show the overall effect Sketch Mechanix has on student understanding of complex concepts. University 1, University 2, and University 5 collecting aspects of each exam. The intention is to follow the growth of these students before and after the Mechanix homework intervention. The collected exam problems were analyzed based on student understanding of the material, with emphasis on the overall score, experimental question scores, and accuracy of FBDs.
3.3 Truss Design Problem Distribution

As part of the study, a new problem type was distributed to participants. This problem focused on the creative application of concepts learned in the class and has been as a “final” in classes. A total of three universities deployed the new problem type. University 1 and University 4 had large courses where students within each section could be randomly split into the Control or Mechanix condition. University 5 had smaller sections that could not be divided into two conditions. Instead, the section was randomly assigned to a condition, with the same instructor consistently teaching all sections. The breakdown of the distribution of the design problem can be seen in Table 4. Students assigned to the Mechanix conditions completed two assignments and the open-ended truss design problem using the Mechanix application at University 1 and University 4. At University 5, students only completed a single assignment before the open-ended truss design problem due to curriculum constraints. Students assigned to the Control conditions used the traditional system employed at the university that was used for the first two assignments and pen-and-paper for the open-ended truss design problem.
Table 4: Initial deployment of the Truss Design Problem was isolated to University 1, with eventual full deployment beginning during the Fall 2020 semester. Two problem types were distributed, the Low Load and the High Load truss problems.

<table>
<thead>
<tr>
<th></th>
<th>Fall 2019\textsuperscript{LL}</th>
<th>Spring 2020\textsuperscript{HL}</th>
<th>Fall 2020\textsuperscript{HL}</th>
<th>Spring 2021\textsuperscript{HL}</th>
</tr>
</thead>
<tbody>
<tr>
<td>University 1\textsuperscript{1}</td>
<td>1 Split Section Control &amp; Mechanix</td>
<td>1 Split Section Control &amp; Mechanix</td>
<td>2 Split Sections Control &amp; Mechanix</td>
<td></td>
</tr>
<tr>
<td>University 2\textsuperscript{2}</td>
<td></td>
<td>2 Split Sections Control &amp; Mechanix</td>
<td>1 Split Section Control &amp; Mechanix</td>
<td></td>
</tr>
<tr>
<td>University 3\textsuperscript{3}</td>
<td></td>
<td></td>
<td>Single Section Mechanix</td>
<td>Single Section Control</td>
</tr>
</tbody>
</table>

LL - Low load problem distributed
HL - High load problem distributed
S - Standard problem statement distributed
A - Aerospace problem statement distributed
I - Imperial system problem statement distributed

The Truss Design Problem prompts users to determine a creative solution to an open-ended problem with set constraints (Figure 3 shows one of the variants). Participants were given a set of criteria for the length and load requirements of a truss and told to design a truss that would meet these criteria. The intention of the problem is to create a solution space that requires students to creatively apply the concepts that they have learned, as solving problems in this manner has been shown to significantly improve student understanding of complex concepts. The design problem is assigned as a mandatory homework assignment, with the grade either being provided based on completion.
(University 5) or based on the accuracy of the max load calculations and meeting the assigned constraints (University 1, University 4). To increase motivation, the student who created the truss that could carry the largest load would receive additional homework extra credit equal to one homework assignment. Participants were also told that they had unlimited attempts to find their proposed solution.

Figure 3: One of the variants used in the study, which was deployed to the participating Aerospace courses taught by the University 2 instructors.

As part of the application's presentation, a class period has been expressly set aside to demo the problem to students, but only to familiarize them with the application and not to give them any guidance on how to best solve the design problem. During this period, participants received a demonstration of the application and received assistance in using the application for the remainder of the class period. This demo and distribution method was altered for online instruction to meet the needs of virtual learning during the COVID-19 pandemic. During the duration of the COVID-19 pandemic (Spring 2020, Fall 2020, and Spring 2021), the demo period was shortened from an entire class period to a 10-15 minute demonstration of the problem UI, although the content of the demonstration was
conserved. In addition to an alteration of the problem demonstration, submissions made by participants assigned to the Control condition could not be turned in by hand. These participants were asked to scan all attempts at the problem and submit them virtually.

Creative application of knowledge is drove the expansion of Mechanix’s capabilities, with participants being asked to design a truss that would meet specific constraints. The constraints include: a set range for the overall length, a maximum length of each member, a maximum permissible internal member load, and a minimum overall load applied to the system. The problem was first distributed within a limited population at University 1, with an easier set of requirements to test the problem and the Mechanix application. Following the problem's pilot distribution, the problem's difficulty was increased by doubling the required minimum overall load applied to the system. This is due to the prevalence of the trivial solution submitted by participants at University 1. The intention in doing this was to increase the challenge and force participants to further explore the design space.

Following the pilot testing of the truss design problem at University 1, Mechanix was modified to track and report each of the individual submissions that students made. Students in the Control condition were asked to submit all attempts they made during the solution process. The wording of the different problems varied depending on the University that it was distributed to, with University 4 and University 5 running variations of the problem first deployed at University 1 starting in the Fall 2020 semester. The problem distributed at University 4 modified the wording to reflect the aerospace class that the problem was distributed to, using an airplane wing truss instead of a guinea pig bridge. The
problem distributed to University 5 used the U.S. customary unit system instead of the metric system as used at University 1 and University 4.

Aside from these changes, the Mechanix application and the associated problem statement were configured the same way, ensuring that students had as similar as possible between the three universities that the problem was deployed. The different variations to the problem statement can be seen in Figure 4 below. One of the modified problem types for the full deployment of the high load problem at University 4 can be seen in Figure 3, which shows the window that student participants saw as part of the Mechanix condition during the completion of the problem. An example of a student submission can be seen in Figure 5. The Control condition was given the same problem on paper, with this handout included in Appendix B.2.

1. Design a truss bridge connected to the pin (left, node A) and roller (right, node B) joints shown.
2. The bridge should span between 10.9 and 13.4 cm between nodes A and B and support a load P of at least 13 N.
3. Each truss member can hold a maximum of 4.5 N in compression and tension.
4. Each node across the top chord of the truss shares the loading equally. For example, if there are 3 nodes across the top of your truss, each node will have an applied load of P/3 downwards.
5. Determine the max value of P before the truss fails.
6. Only draw the truss and applied load.
7. All distances you input are in cm. All angles are in degrees.

Figure 4: The different variations in the distributed problem statements can be seen above. Not pictured is the variation used at Univ. 3, which altered the problem to use the U.S. customary units.
Figure 5: Example of a drawn solution. The truss design problem asks students to draw the proposed truss, followed by the input of the proposed geometry and then finishes with the application of the loads.

Participants assigned to the Mechanix condition were provided access to the online portal for the problem and asked to complete the problem there. This portal provided the problem statement, a field to draw their proposed solution, an answer field for the maximum permissible load, and provided feedback if the provided solution was incorrect. The application required students to draw the FBD of their proposed solution, enter the dimensions, and then finalize the design with the load(s) applied to the truss. The types of feedback provided include: minimum load not achieved, minimum length not achieved, truss fails under calculated load, invalid truss geometry, or a wrapper error message should the submissions have an unforeseen cause of failure. The Control condition received the problem statement on paper and was asked to turn in all of their attempts, receiving feedback if they went to instructor office hours or at the end of the grading period.
Participants in both conditions are informed that they are allowed an unlimited number of submissions, with the hope that all students will explore the solution space.

3.4 Control Participant Submission Coding

Submissions made by participants at University 4 assigned to the Control conditions were coded based on feedback that was formatted similar to feedback provided by the Mechanix application. The addition of a few types of feedback that Mechanix automatically catches were added due to more errors being possible in paper submissions than through the Mechanix application. The feedback was coded by two members of the research team, with an Inter-rater agreement being conducted to ensure the type of feedback provided was consistent.

Coded feedback was based on the most likely cause of error made during the solution process. Mistakes in the creation of the FBD were assumed to be the root cause, as these mistakes trickled down into the solution attempt through incorrect assumptions, erroneous force vectors, or through miss-placed truss members. This coding methodology ensured that submissions that could receive multiple types of feedback received the most likely cause of failure. For instance, if a participant drew a truss with more than two supports, they more than likely make both an error in their calculations due to the inclusion of extra force vectors, as well as making a major conceptual error related to the generation of the FBD. In this situation, the participant would receive the “Conceptual Errors Made” feedback. The types of coding feedback and an associated example can be seen in Table 5.
Table 5: The following table details the types of feedback given to participants assigned to the Control conditions. Some of these types of feedback are impossible to receive through the Mechanix application

<table>
<thead>
<tr>
<th>TYPE OF FEEDBACK</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculated Truss Fails Under Load:</td>
<td><img src="image" alt="Example Image" /></td>
</tr>
<tr>
<td>This was assigned when a proposed truss max load causes one or truss member to fail. Mechanix reports this as a possible error.</td>
<td><img src="image" alt="Example Image" /></td>
</tr>
<tr>
<td>Minimum Length Not Achieved:</td>
<td><img src="image" alt="Example Image" /></td>
</tr>
<tr>
<td>It is assigned when the student makes an error where the proposed truss solution does not meet the minimum length requirement. Mechanix directly reports this feedback to the student.</td>
<td><img src="image" alt="Example Image" /></td>
</tr>
<tr>
<td>Conceptual Errors Made:</td>
<td><img src="image" alt="Example Image" /></td>
</tr>
<tr>
<td>Refers to when a student makes a significant conceptual error in the calculation of the truss. Used when a participant drew a truss as a frame, made major conceptual errors in calculations or representation, or misunderstood truss solution methods.</td>
<td><img src="image" alt="Example Image" /></td>
</tr>
<tr>
<td>Incorrect Application of Load:</td>
<td><img src="image" alt="Example Image" /></td>
</tr>
<tr>
<td>This is assigned when a conceptual mistake is made by the student, usually associated with thinking of the system as a beam or as a frame. This situation is not possible in Mechanix.</td>
<td><img src="image" alt="Example Image" /></td>
</tr>
<tr>
<td>Invalid Geometry:</td>
<td><img src="image" alt="Example Image" /></td>
</tr>
<tr>
<td>Refers to the creation of an invalid Planar truss, either because the submitted truss has members that cross, invalid support placement, or due to an over-constrained truss</td>
<td><img src="image" alt="Example Image" /></td>
</tr>
<tr>
<td>Unstable Truss:</td>
<td><img src="image" alt="Example Image" /></td>
</tr>
<tr>
<td>Refers to the creation of an under-constrained system, in which the system would collapse under a new lateral load. Mechanix does not even recognize this as a valid truss geometry.</td>
<td><img src="image" alt="Example Image" /></td>
</tr>
</tbody>
</table>
3.5 Pilot Study: Dynamics Distribution

As part of the effort to bolster the capabilities of Sketch Mechanix, the AI algorithm was expanded to include architecture capable of distributing Dynamics problems, which was tested at a large public university in California (University 3). This university was selected based on the diverse student body found at the school, as it would enrich the study’s results. A problem set was developed using questions that had previously been distributed at the school, with one of the example problems visible in Figure 6 below.

The distribution of Dynamics problems at the participating university is included in Table 1. Due to smaller class sizes at the participating university, the experiment conditions were assigned to class sections instead of on an individual basis. Participating class sections were assigned with the intention of creating an equal number of Control condition (pen and paper homework method) participants and Mechanix participants over an equal number of semesters. This homework was graded based on the accuracy of the final solution, as well as the overall accuracy of the drawn FBD provided as part of the problem. A single Dynamics problem set was distributed in this manner, with the remainder of homework problems using the traditional homework method of participating university.

Student participants at the University 3 followed the same experimental design as other schools, with slight deviations due to the type of class that the problems were deployed at, assigning an entire class section to either the Control or Mechanix condition. Student participants completed pre- and post- diagnostics in the form of Concept Inventories, although participants at University 3 completed the aDCI instead of the FCI.
and SCI. The aDCI is distributed in the same manner as the other diagnostics and was the major focus of the Dynamics Pilot study at University 3, comparing the pre- and post-diagnostics to determine the difference in learning outcomes through the use of the Mechanix application when compared to a traditional homework method.

Figure 6: One of the Mechanix problems developed which uses Dynamics concepts, which was deployed to Dynamics classes at University 3.

In addition, an experimentally controlled exam question was placed on the midterm exam following the Mechanix intervention. This exam question was selected specifically to test knowledge that is thought to be developed by the Mechanix application, with comparisons being made between the sections that had different experimental conditions assigned to them. This comparison followed the comparisons made at other schools, analyzing the overall efficacy of the proposed FBD, the accuracy of the proposed answer to the experimental question, and analyzing the overall score between the two conditions.
CHAPTER 4. RESULTS

4.1 Mechanix Homework Distribution Results

The results of the study on the efficacy of the Mechanix application are broken down by the summary of the distribution of each of the measures. This is organized for the clarity of the impact of said instrument, with descriptions of the results collected at each of the universities being summarized and discussed for each of the instruments. The results of the “pilot” studies are further broken down, with each having its own dedicated results section.

Additional homework results were collected as part of the study; however, due to existing conditions at the universities at different stages of the study, both a lack of participation and a known self-selection bias have significantly decreased the participating population available for analysis. Existing conditions have included: transitions in instructors due to paternity leave and sickness, disruptions due to wildfires during multiple semesters, interruptions due to an unexpected blizzard, and distractions due to the ongoing COVID-19 pandemic. Due to all of these disruptions, some of the collected data has been deemed “unusable” due to the biases induced in the data.
4.1.1 FBD Problem Set Scores

A general bodies FBD assignment was the first Mechanix homework intervention of multiple Mechanix homework assignments in a semester. The results of each homework assignment between the two conditions were compared, with the overall results of each section showing no variation between the Control and Mechanix conditions on the first FBD homework assignment. The results from this first homework assignment can be seen in Figure 7 below. There was no statistically significant difference at any of the universities, with almost no difference between the Control and Mechanix conditions.

Due to this trend, focus groups were conducted at University 1 during the Fall 2019 semester and at University 5 during the Spring 2021 semester. Conducted by either members of the research team (Fall 2019) or by a group of outside evaluators related to the study (Spring 2021), these focus groups were asked a series of questions, with the recorded results coded based on the similarity of the responses. An overwhelming number of focus group participants agreed that there was a general frustration related to learning a new application halfway through the semester, as well as an agreement that the new application had an associated learning curve. Due to the focus group responses during the Fall 2019 semester, deployment of the application in the following semesters attempted to demo the application in such a way to reduce new user anxiety. This was done through the introduction of tutorial videos, the inclusion of practice assignments, and the expansion of a Mechanix instruction hand out.
Figure 7: The data summarized above is taken from Fall 2019, Spring 2020, and Fall 2020. All schools in the participating data sets ran the general bodies assignment as the first Mechanix activity, with no statistical difference between the Control or Mechanix conditions.

Figure 7 shows the collected and analyzed results of the FBD homework assignment, showing that the additional emphasis on problem demoing may not have overcome the issues related to the learning curve. At all participating Universities that deployed the FBD homework assignment as part of the experiment, results exist as a trend in favor of the Control condition on the first assignment. Following the Fall 2019 assignment, participants assigned to the Mechanix condition were provided a not for credit practice problem set of tutorial questions, as to help the participants understand the application. Of the participants that was assigned to the Mechanix condition, around 15% of participants even attempted this problem during the Spring 2020 semester and onward. Interestingly, participants that completed this practice assignment scored marginally higher than Mechanix participants that did not, however to a minor degree.
4.1.2 Mechanix Assignment Truss Problem Set

A method of joints truss analysis assignment was deployed as the second Mechanix intervention at three of the five universities, with the University 5 distributing the problem set as the first Mechanix intervention. After controlling for differences in grading between the two homework systems, the effects of persistent usage of the Mechanix system presented positive results. Figure 8, below, shows results from sections deploying method of joints truss analysis assignment. Schools that deployed the truss assignment as a second intervention saw statistically identical homework scores (University 2) or a difference in favor of the Mechanix condition. Results from University 1 and University 4 were compared using a t-test assuming unequal variance, which marginal statistical significance at University 1 ($t(37)=2.002$, $p=.106$) and strong statistical difference at University 4 ($t(186)=3.71$, $p<.001$).

Figure 8: Comparison of the data collected at the universities during Fall 2019, Spring 2020, and fall 2020. Results suggest continued usage of the application assist student performance using Mechanix.
Due to curriculum constraints, University 5 deployed Mechanix as a single intervention through the truss assignment. Results from this university initially differed solely based on the method of grading conducted, with differences in the grading of the submissions made by the Control condition skewing the results against the Mechanix condition with marginal statistical significance in favor of the Control condition \((t(56)=1.69, p=.095)\) (Figure 9). These results were consistent across all semesters of distribution and are similar to the results of the FBD problem set distribution at all other participating Universities.

![University 5 Truss Assignment](image)

**Figure 9:** Results of the Mechanix intervention at University 5 suggests that the usage of the Mechanix application has a learning curve.
4.2 Exam Results

In addition to the homework problem sets collected as part of the study, specific exam problems were assigned to gauge student understanding of the material reinforced by the Mechanix intervention. The exam problems focused on the sketching and usage of FBDs for the completion of a Truss problem, a skill that is thought to be improved by the Mechanix application. The growth of student achievement on exams was measured during the study by collecting multiple exam questions at specific universities (University 1 and University 2). At all universities, exam results were collected following the Mechanix intervention, specifically to determine the effect of the Mechanix application on student performance. Results included the overall exam grade, the student performance on specific exam questions, and finally, the accuracy of FBD created.

4.2.1 University 1 Exam Results

Exam results from University 1 were compared, with a difference appearing in favor of the Mechanix condition following the Mechanix intervention. This occurrence can be seen below in Figure 10. Following the Mechanix homework intervention, Mechanix participants had a slightly better performance their their peers. This can be seen through a trend in favor of the Mechanix condition on Exam 2, which disappears by the time of the final exam. On the experimentally controlled exam question, this difference exists as a statistically significant difference, in favor of the Mechanix condition ($t(34)=1.81, P=.079$), suggesting the usage of the Mechanix application had a positive influence on student understanding related to truss concepts. These results can be seen pictorially in Figure 11 below, which reflects data collected during the Fall 2019 semester.
Collection of future exam data sets at University 1 were interrupted by the COVID-19 pandemic and were deemed unusable. These interruptions started during the Spring 2020 semester, during which exams were not distributed in a secure environment. This was due to the sudden shift in instruction related to the health requirements associated with the COVID-19 pandemic, and resulted in evidence of “foul play” in the exam results. Additional interruptions occurred during the Spring 2021 semester, during which a drastic shift in the instructor caused differences in instruction, which past research have an influence in differences in the exam results (detailed in Pilot Study: Dynamics Problem).

Figure 10: On all three exams, both the Mechanix and Control conditions performed equally. A trend occurred following the Mechanix intervention (Exam 2), suggesting the Mechanix application may have impacted student understanding.
Figure 11: Statistical difference between the conditions on the experimentally controlled question. While students did not show a statistically higher grade on the exam overall, this result suggests that the application improve the Mechanix condition scores.

4.2.2 University 2 Exam Results

University 2’s exam results were less impacted by the pandemic, although University 2 had smaller class sizes reduced the collected sample size. At University 2, the condition assignment was given to an entire class section and as such, comparisons were made between different semesters, only deviating during the final semester of data collection (Fall 2020) as to keep the data sets populations as similar as possible. Under these conditions, the Mechanix intervention occurs before Exam 2, with no statistical difference between the Control and Mechanix conditions across any of the exams. These results can be seen summarized in Figure 12 shown below, with the results being collected and analyzed across multiple semesters, indicating consistency across all data collected at University 2. The implications of these results are discussed in-depth as part of Error! Reference source not found. discussion section.
Figure 12: The overall exam scores at University 2 (collected from Fall 2018, Fall 2019, and Fall 2020) reflect the other instruments distributed at University 2, with no indication that the Mechanix application harmed student understanding of the material.

4.2.3 University 5 Exam Results

Additional exam results were collected from University 5, with the results from the Fall 2018 and Fall 2019 semesters being analyzed in a similar fashion as the other two schools. The resulting analysis had some confounding factors that restricted the overall efficacy of the instrument, specifically the impact of the time of day on student achievement. The two semesters that were analyzed had conducted two separate sections which saw the distribution of the Mechanix application. During the Fall 2018 semester, the AM section was assigned to the Control condition, and the PM section was assigned to the Mechanix condition. To contrast this, during the Fall 2019 semester, the section assignments were flipped. The results on the experimental question are summarized below in Figure 13.
Figure 13: Through the analysis of the exam scores at University 5, it was discovered that the time of day had a larger impact on student performance than the introduction of the Mechanix application.

A T-test assuming unequal variance was conducted within each semester and comparing the comparison of the time of day, which yielded interesting results. Within a given semester, the PM section performed better than the AM section, with statistical significance in favor of the Control condition during the Fall 2018 semester (t(91)=1.91, P-value=.058) and a trend in favor of the Mechanix condition during the Fall 2019 semester. By comparing the results by controlling for the time of day, it was discovered that the results become more in line with the other schools when the AM or PM of either section are compared, with a trend in favor of the Mechanix condition during the AM sections, and a trend in favor of the Control condition during the PM section.
4.3 Concept Inventories

Concept inventories were deployed as pre- and post-semester measurements for the project. The concept inventories deployed previously developed, peer-reviewed instruments to discern if the Mechanix application directly impacted student conceptual understanding. These peer-reviewed instruments were also used to determine any systemic differences concerning the prerequisite knowledge of participants within the associated classes. As stated before, peer-reviewed instruments were used in the deployment of the Mechanix application: the FCI, the SCI, and the aDCI.

4.3.1 Force Concepts Inventory

Almost all participants who completed the FCI showed no significant difference between pre- and post- distributions of the CI, regardless of the assigned condition. This trend can be seen in Figure 14, shown below, which presents the difference between Control and Mechanix students on the CI. It should be noted that while there appears to be no difference between the two conditions at a single university, some of the schools performed worse than other schools, which is believed to be due to the participants at the participating University. At Universities 1, 2, and 4, participants are almost universally enrolled in an engineering major, while participants at University 5 are enrolled in Construction management, and therefore the lower scores of non-engineering participants is not surprising. Regardless of the micro difference between the schools, the overall results of the FCI indicated that students had a statistically equal conceptual understanding of concepts covered by the FCI through the entire class, and the Mechanix application had no effect on students’ knowledge of these concepts.
4.3.2 Statics Concepts Inventory

The analysis of SCI showed growth in student performance from pre- to post-examination in both Mechanix and Control conditions, with no significant difference across the conditions. The exact difference in the growth is mildly dependent on the school from which the data is collected. The data from the Fall 2019 semester represents the overall results throughout the study. These results can be seen in Figure 15 below, showing each condition's growth from pre- to post- semester distribution of the SCI. One specific thing to note, the introduction of the Mechanix intervention has no statistically significant impact on student understanding of the concepts on the inventory. Analysis was conducted based on the overall scores, as well as the subscore of the SCI.
Figure 15: Like with the FCI, the pre- and post-semester distribution of the SCI showed no significant difference between the Mechanix and Control homework conditions at all participating schools.

The analysis of the Fall 2019’s subscore at University 1, seen in Figure 16, showed that specific sub-categories consistently have a lower score than the other categories. This difference was especially prevalent on concepts that had little relationship to concepts reinforced by the Mechanix application. Because of these differences, as well as the evidence of low effort among some responders, the subscore results were analyzed to determine which subscores were perceived by participants as easier. This analysis determined that three of the subcategories were answered correctly at a higher rate than the other six subcategories, suggesting that these question sets are easier than others. Because of these differences, a reordered SCI was distributed to participants following the Fall 2019 results, with the easier questions front-loaded on the distributed reordered survey.
Figure 16: These results show that certain questions are more commonly answered correctly than others, possibly suggesting that certain questions are easier.

The Reordered SCI survey was distributed starting during Spring 2020, which coincided with the beginning of the COVID-19 pandemic. The corresponding shift to online learning decreased the participating population in the study. Despite the shift in education and the decrease in participation, the reordered survey was deployed to multiple schools during this semester. No significant results were discovered between the initial deployment and the reordered deployment, with the only exception being at University 5. This deployment saw significance in favor of the reordered survey as calculated through a two-tail T-test assuming unequal variance ($t(63) = 1.99, p = 0.049$). Due to the mixed results directly related to the lack of harm that the reordered survey produced, as well as the small participating population, the reordered survey participation did not yield enough results, and further deployment was necessary. The reordered SCI was distributed again during the Fall 2020 semester, with classroom instruction more similar to a traditional setting at University 4 and University 5. A comparison of the results from Fall 2019 to Fall 2020 can be seen in Figure 17 below.
Following the reorder of the SCI, all had marginal improvements, with the results continuing to reflect previous results, except at University 5.

Results collected from University 5 were compared using a T-test assuming unequal variance, which yielded a statistically significant difference in favor of the reordered data set ($t(91) = 3.11$, $p$-value<.001), indicating the participants assigned to the Mechanix condition who completed the SCI during the Fall 2020 semester outperformed their peers that completed the Original SCI distributed prior to the Spring 2020. Further analysis was conducted with a comparison of the Spring 2020 data set and the combined Fall 2020 and Spring 2021. A T-test assuming unequal variance showed growth in both conditions; however, the growth in the Mechanix condition is more significant than the other comparison. Analysis comparing the Post SCI Mechanix conditions through a t-test assuming unequal variance confirms the significance of this data point ($t(91)=3.11$, $p=.0024$), in favor of the reordered SCI survey, and can be seen in Figure 18, with further significance shown in favor of the Mechanix group.
Figure 18: Through comparing the original SCI results and the reorder SCI, it was discovered that both the Control and Mechanix participants performed better on the reorder survey.

4.4 Pilot Study: Dynamics Problem

As part of the impact of academic efficacy of the Dynamics Mechanix problem set, exam questions were collected over the course of three semesters at University 3. Due to the smaller student body enrolled in the participating class sections at University 3, the participating sections were all assigned to one condition. The results of these class sections were compared across the semesters, with the collection of these results being interrupted by multiple factors. The distribution of the homework assignment assigned as part of the experiment was disrupted by a combination of natural disasters and environmental factors.

Data collection was severely disrupted by wildfires during the Fall 2018 and Fall 2019 semesters, either completing disrupting data collection due to canceled classes (Fall 2018) or delaying and damaging the expected outcomes (Fall 2019). Additionally, inconsistencies occurred in instruction and grading practices during the Fall 2019 semester,
as the original instructor went on family leave part way through the semester. These compounding factors made the results from the Sketch Mechanix Dynamics problem sets inconclusive, with the research team electing to leave portions of the collected data unanalyzed or under-analyzed due to the time constraints on the study. Part of the data was analyzed through a combination of efforts made by members of the research team, with the exam results and aDCI being the only aspects fully analyzed.

Initial coding results from the collected Dynamics exams indicated that differences in instructors during the Fall 2018 and Fall 2019 semesters might have had an unforeseen impact on the comparibility of the data sets. Both the composition of the overall exam, as well as the wording and weighting of the experimental exam question, changed from the Fall 2018 semester to the Fall 2019 semester. The two exams distributed an similar, but not identical exam question, although the Fall 2019 exam put less emphasis on the question, reducing the overall point value. Changes in the wording also impacted the results of the question, as the Fall 2018 exam required the sketching of the associated FBD, while the Fall 2019 exam did not. These reasons, as well as a change in, made the results of the Mechanix intervention inconclusive.

These differences were analyzed, with the decision being made that in order for the results to be comparable, one or both of the data sets needed to be regraded. The Fall 2019 data set was selected for regrading using the Fall 2018 grading rubric, as this rubric was written by the original instructor of the Dynamics sections. This effort was conducted as part of a for-credit project by a member of the research team, yielding the results shown in Figure 19 below, which contains both the regraded exam question as well as the rescored exam total. While the regarded results were made more comparable than the previously
analyzed data set, the pre-existing differences in exam distribution indicated that even with the data sets being graded in an identical way the results might not be indicative of the impact of Mechanix on student achievement.

![University 3 Semester Comparison]

**Figure 19:** The results from the regrading effort and resulting analysis indicates that the administered exams during the Fall 2018 and Fall 2019 semesters may not be indicative of the learning outcomes of the Mechanix participants

The regrading efforts produced a richer data set, as the question that was distributed on the Fall 2018 midterm exam had higher weighting, and therefore had larger differences in the answers submitted by students. The results of this regrading effort and resulted in new statistically significant data points, as proven by t-tests assuming unequal variance. The first being that the students of the Fall 2019 semester (Mechanix intervention) outperformed their peers of the Fall 2018 semester (Control condition) on the overall exam \((t(46)=2.59, \ p\text{-value}=.015)\). The second aspect of the data was that the students assigned to the Control condition outperformed their peers assigned to the Mechanix condition on the experimental test question \((t(47)=3.7, \ p\text{-value}<.001)\). The expectation is that students that used the Mechanix application would perform better on all aspects of the exam.
especially the experimentally placed question. The difference in exam score in favor of the Mechanix group on the overall exam, while against the Mechanix group on the experimentally placed question shows that the differences in exam distribution may make the results from the two semesters uncomparable.

An additional data point on the exams was considered during this regrade effort, that being the rate to which students accurately drew a FBD associated with the experimental question. Drawing the FBD was a required aspect of the exam distributed during the Fall 2018 (Control) semester, while it was optional during the Fall 2019 (Mechanix) semester. This simple difference caused this analysis to be flawed from the beginning, as it has been theorized the inclusion of the drawing in the exam score caused more students to draw the FBD in the Control condition than in the Mechanix condition, as the participants that were assigned to the Control condition created more accurate FBD’s to a significant degree (t(51)=2.62, P-value=.011). These discrepancies in exam results, as well as the conditions of each semester of data collection, indicated that the results of the data collected at Univ. 3 may be incomparable, with the analysis of the homework assignment not being conducted, due to the exam results and the experimental inconsistencies.

The aDCI was distributed only at University 3, which replaced the standard distribution of the FCI and SCI. Participants completed the aDCI as part of the set of Pre- and Post- diagnostics, with comparisons being made between the sections assigned to Control and Mechanix conditions. The aDCI asks students to complete 11 Dynamics-related questions, as opposed to the 27 questions distributed in the full DCI survey. The results of this instrument (Figure 20) seem to be impacted by the conditions at Univ. 3 that
occurred during the periods of data collection, with results being significantly lower than expected. As seen in Figure 20, the results indicated that the usage of the Mechanix application did not have an impact on student understanding, with participants in both experimental conditions experiencing equal growth throughout the semester. This indicates that the Mechanix application had no effect on student understanding of the concepts tested by the aDCI.

![aDCI Condition Outcome Comparison](image)

**Figure 20:** The Mechanix intervention had no impact on the results on the aDCI. This is consistent with Concept Inventory scores at other universities, although the conditions during data collection are expected to have affected the results.

The lower scores on the aDCI may indicate that the concepts tested on the Dynamics Concepts Inventory may not be representative of the understanding of the surveyed students, or consistent with the topics taught in class. These results are expected to be heavily impacted by the unexpected conditions at University 3. The fall 2019 semester was directly impacted by a shift in instructors, and the spring 2020 semester being impacted by the shift in education due to the COVID-19 pandemic. Regardless of the reason, the results of the aDCI collected during these semesters should be carefully
considered, as compounding factors may make the data sets unrepresentative of the overall learning outcomes due to the Mechanix intervention at University 3.

4.5 **Truss Design Problem Results**

4.5.1 *Truss Design Problem Pilot Study*

The Truss Design Problem has been deployed from Fall 2019 to the present, under different environmental conditions and with varying rates of participation. Deployment began in the Fall 2019 semester at University 1, with a simple problem intended to test the AI and the problem statement. The participating sections at University 1 for all semesters were divided into split Control/Mechanix settings to decrease the effect of the difference in instruction across sections of the class. For this initial testing, a trivial solution made up of equilateral triangles was known to exist, with the assumption that the participants would not immediately find this solution and would iterate through multiple attempts. The initial hope was that student participants would attempt to find a stronger solution based on the available extra credit that would go to the student that created a solution that would hold the largest load. This trivial solution can be seen below in Figure 21.
Figure 21: The figure above shows the trivial solution of the low-load design problem.

Results from the pilot distribution did not align with these expectations. The low load problem statement asked participants to produce a truss that could hold a minimum of 7.5 N. Participants of the two conditions produced mean truss load of 8.59 N and 7.98 N for the Control (n=19) and Mechanix (n=18) conditions, respectively. A two-tail T-test showed that there was no statistically significant difference between the means of the two conditions (t(35) = 0.13, p-value=.895). These results can be seen graphically in Figure 22. Between the two conditions, 30 students submitted the exact same answer, which coincided with the trivial solution. From this initial deployment, it was decided that a more difficult problem needed to be deployed to increase the challenge of the problem and the corresponding richness of the data. It was also determined that the new problem type was stable enough for wider deployment, as a majority of the participants saw no significant software bugs during the process.
By increasing the required load applied to the truss system in Spring 2020, it was assumed that the student participants at University 1 would explore more of the design space. Additionally, the application’s grading capability was improved by tracking the total number of attempts each student submitted in order to understand each student’s exploration of the solution space. COVID made it difficult to test this theory, as the Truss Design problem homework was assigned the week University 1 switched to online instruction. This caused a disruption in the assignment, resulting in lower participation than the initial deployment during the Fall 2019 semester. To comply with the health guidelines, the due date for the problem was extended, and students in the Control condition were informed to scan and upload their solutions; students in the Mechanix condition were tasked with finishing the assignment using Mechanix, and the grading process for the entire class was also made more lenient. Less emphasis was put on asking participants to submit all of their attempts, specifically to reduce the load on students during the period.
Of those that submitted answers during the COVID transition period, the changes to the problem had an impact on the diversity of submissions as almost every student submitted a different solution to the problem statement. However, the overall participation was lower due to the tumultuous period that the problem was assigned. This decrease in participation was not equal between the Control and Mechanix conditions, with more submissions occurring from the Mechanix condition. 29 of the 52 participants submitted an answer. 12 of 27 Control participants submitted a final answer, and 17 of 25 Mechanix participants submitting a final answer. To determine if there was a difference in submissions made by the two conditions, a Chi-square test was conducted. It was determined that the difference in the rate of submissions made by the two conditions has no significant difference, but the trend does exist between the two conditions (p-value=.36).

The prevailing theory concerning the lack of participation is a self-selection bias existed due to the difficulty in submitting the Control results, with participants assigned to the Mechanix condition having a more convenient method of submitting a final result that they had previously used. Mechanix may have also been more motivating as the instant feedback provided by Mechanix may have also encouraged students to complete the assignment. Of the answers submitted, an almost equal percentage of students submitted an answer that was taken as being correct, with the overall submission results being seen in Figure 23.
Figure 23: Overall participation during the shift to online instruction caused a general lack in participation in the Truss Design Problem.

Following the pilot testing of the Truss Design Problem during the Fall 2019 semester, Mechanix was modified to track and report each of the individual submissions that students made during the assignment period, while students in the Control condition were asked to submit all attempts that they made during the solution process. Through this new method of tracking the solution process, it was discovered that there was a significant difference between the Mechanix and Control conditions during the Spring 2020 semester, in favor of the Mechanix condition. This difference was confirmed through a two-tail t-test assuming unequal variance (t-value=3.47, dof=14, p=.003) and can be seen in Figure 24 below.

It is assumed that the main reason for this difference is due to the lack of work submitted by Control participants, as most members of the Control condition only submitted their final answer. From looking through the submissions made by the Mechanix condition, it is apparent that Mechanix at a minimum provides an easier method of tracking
student solution attempts than a paper-based homework method. It can be extrapolated through the number of different submissions made by students that the increase in the challenge created a situation where the instant feedback of the application encouraged the students to explore more of the solution space. This feedback is one of the main differences between the Control and Mechanix condition. A participant in the Mechanix condition will automatically know if a mistake is present in their proposed solution, while the Control participants will only know if a mistake has been made once the grading process has been concluded or through interaction with the instructor.

![Figure 24](image)

**Figure 24:** A comparison of the total number of the mean total answers submitted per user on the Truss Design mode, showing that the Experimental condition submitted a statistically different number of submissions than the Control condition.

As a result of COVID-19 altering the sample size of the truss design participation, the impact of the application on the exploration of the sample space has been diminished. The total number of students in the Control and Mechanix conditions who got an accurate solution to the problem was far smaller than the total number of students who agreed to participate in the study, with only 19 participants calculating an accurate maximum load for their proposed design (Figure 25). Within this small analyzed group of participants, no
significant difference existed between the mean load of the two experimental conditions, although there was a trend towards the Mechanix condition performing slightly better than the Control condition ($t(13)=0.32$, $p=.76$). With this initial data analyzed, it was expected that with a larger sample size, the participants that completed the assignment using the Mechanix application would produce proposed solutions that could hold a larger load than those who used a traditional pen-and-paper solution method.

![Spring 2020 Uni 1 Creative Design Mean Truss Load](image)

**Figure 25:** A comparison between the mean calculated load the submitted trusses could handle, showing that the spring 2020 Variation of the problem was solvable and that the experimental condition trended towards having better results.

### 4.5.2 Expanded Problem Distribution

For the full distribution of the Truss Design problem at University 1, of the students that participated in the study in a meaningful way, the maximum load that the truss could carry was analyzed. Of these participants, the Control Condition achieved an average of 14.68 N, with the Mechanix Condition achieving an average of 14.64 N on their final truss submissions. A t-test assuming unequal variance was conducted ($t(37)=0.048$, $p$-
value=.96), showing no significant difference between the two conditions. These results can be seen in Figure 26. The results Figure 26 are comprised of three classroom sets, two conducted in Spring 2021 and one collected in Spring 2020. Of note, the Spring 2020 data set was significantly impacted by the Covid-19 pandemic as it was distributed during the same period that instruction methods were shifted at University 1. This data was condensed due to the small participating populations during the semesters of data collection.

![University 1 Truss Design Mean Load Spring 20 & Spring 21](image)

**Figure 26: The results at University 1 seem to suggest that in high-stress situations, the Mechanix application does not harm student performance.**

In comparison to the Truss Design Problem distribution at University 1, University 4 saw a more negligible impact of COVID-19 during the expanded problem distribution. This analysis was considered especially important, as University 4 saw the largest participation rate across all deployments of the problem. Unfortunately, the rate at which participants produced complete responses differed between the two conditions, with 24 Control participants and 54 Mechanix participants out of 87 Control participants and 91 Mechanix participants. The difference in correct submissions and their associated maximum proposed load shows the influence of the scaffolded feedback built into
Mechanix. With participants assigned to the Control participants submitting a final truss that on average held 14.57 N while the Mechanix participants achieved an average of 15.94 N. This difference was analyzed using a t-test assuming unequal variance, with Mechanix participants producing a statistically significant higher mean load \( t(85)=2.27 \ p-value=.028 \) (Figure 27). These results partially confirm the hypothesis that Mechanix provides students with the opportunity to explore more of the open-ended problem’s design space, as those assigned to the Mechanix condition produced stronger Truss designs than those assigned to the Control condition.

Figure 27: The mean calculated load shows that the difference in the experimental conditions is in favor of the Mechanix condition.

Further analysis was conducted using the results of University 5, with students there being awarded credit based on participation and not accuracy. These results were analyzed similarly, although trends between the two semesters of distribution suggest that the design problem may have been too complex for the students, as only around 5% of the participants found a working solution. The results also indicated that students who submitted a correct final answer might have found this final solution through too much collaboration. The only
solutions that met the problem criteria were nearly identical to the solution proposed by the Instructor at University 5, as a few participants discussed this solution during office hours with the professor. For this reason, the overall load results at University 5 was not fully analyzed, although other aspects of the collected data were analyzed similarly to University 1 and University 4.

4.5.3 Comparison of the Provided Feedback

Submissions made by participants assigned to the Control condition at University 2 were coded according to Table 5, with the process being repeated by two research members. An Inter-rater agreement of 64% was achieved, with differences in the Inter-rater agreement came from discrepancies concerning what caused the participant to make a mistake in their solution process. After discussing these discrepancies, clarifications were made in the codebook to avoid future discrepancies in this coding and an Inter-rater agreement of 85% was achieved. The corresponding results were grouped based on the most likely cause of error in the given solutions. The coded Control condition feedback was then compared to feedback automatically given to Mechanix participants, with results suggesting the Mechanix application supported idea generation through the scaffolded feedback provided as part of the application.

The lack of automated feedback given to students in the Control condition is probably the reason behind the higher rate of errors made in final submissions. At University 4, the participants assigned to the Control condition made errors related to the application of load, incorrect assumptions regarding support placement, or incorrect creation of truss geometry, while members of the Mechanix condition mostly submitted
trusses that had accurate max loads. This difference is likely because the Mechanix application provided instant feedback on most of these mistakes. Of submissions made by Control participants, only 30% were graded as correct. In comparison, only 40% of submissions made by Mechanix participants submitted an incorrect answer. As a check to determine if the difference in submissions was substantial between the traditional and Mechanix conditions, the rate of correct submissions of the two conditions was compared using a Chi-square test. The Chi-square test results showed statistical significance (p-value=.013) in favor of the Mechanix participants, meaning the difference in submissions is not due to random error. A total of 23 participants assigned to the Control condition submitted a correct final truss that met all design requirements, while a total of 54 participants assigned to the Mechanix conditions submitted a correct final truss that met all design requirements.

Of all submissions made by the Control condition at University 4, 29 submissions were made that would have been corrected automatically by Mechanix, with common errors mostly related to invalid geometry, incorrect application of load, or incorrect assumptions made (Figure 28). By comparison, the errors made by members of the Mechanix condition are due to participants either making errors in calculations or failing to meet one of the design requirements, with over half of the errors made receiving feedback that the submitted truss did not meet the stated design requirements of the problem statement. Additionally, the only feedback type in which more Mechanix participants made errors was related to submitting a final truss that fails under the provided max load, meaning their proposed truss did not meet the minimum load requirements for the design problem.
Figure 28: The breakdown of the errors made by students by condition shows that the Mechanix application reduced both conceptual errors and the overall rate of mistakes made through the Mechanix application.

4.5.4 Condition Submission Count Comparison

The total number of submitted answers each student submitted was also recorded and analyzed. This analysis seems to confirm that students at all participating universities attempted a wide variety of solutions when using the Mechanix application. Due to the pilot study results, participants assigned to the Control conditions were reminded multiple times to submit all of their attempts to the problem. As the overall participant population at University 4 was significantly higher than previous deployments of the problem, it was expected that a more significant subset of Control participants would submit more than only the final submission.
As shown in Figure 29, this was not the case. For the Control condition, only submissions with accompanying work were considered in this analysis, meaning if a student simply sketched a truss it was not considered an attempt. The mean number of answers submitted was 1.10 for the Control and 8.72 for the Mechanix condition, with the resulting t-test analysis showed that the differences were statistically significant ($t(73)=6.54$, $p<0.001$). This difference confirms that the Mechanix application either captured more attempts made by the student participants or that the Mechanix application allowed participants to produce more submissions and explore more of the solution space. However, the lack of Control condition submissions does not mean that the difference in student submissions cannot solely be attributed to the exploration of the solution space. The total number of submissions made by participants at University 5 (summarized in Table 6) are consistent with the results collected at the other participating universities.

Table 6, Breakdown of Student Submissions by Condition at University 5.

<table>
<thead>
<tr>
<th></th>
<th>Fall 2020 (n = 20)</th>
<th>Spring 2021 (n=33)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>N/A</td>
<td>1.05</td>
</tr>
<tr>
<td>Mechanix</td>
<td>8.11</td>
<td>17.5</td>
</tr>
</tbody>
</table>
Figure 29: Shown in one figure, both the University 4 and University 5 experimental conditions are consistent with previous submission counts seen in the pilot distribution at University 1.

4.5.5 Comparison of the Trivial Solution

Through the deployment of the problem set at University 4 during the Fall 2020 and Spring 2021 semesters, it was discovered that the previously existing trivial solution could be slightly modified and work for the more difficult problem specifications. Despite this similarity, less than 10% of all participants submitted the modified trivial solution as their final solution. This is far less percent than the pilot distribution, as close to 80% of students in the pilot distribution submitted only the trivial solution. This difference suggests that the improved problem statement had an impact on the exploration of the design space, as only a small minority found the simple solution to this complex problem. This new trivial solution, shown in Figure 30, requires no more calculations than the previous one.
Figure 30: By Mirroring the low load trivial solution about the X-axis, a solution can be achieved that exceeds the high load problem statement. This new trivial solution was not submitted at the same rate as the low-load trivial solution.
CHAPTER 5. DISCUSSION

5.1 Mechanix Has a Learning Curve

Through the continued deployment of the Mechanix application at the participating universities, it was discovered that the application seemed to have a learning curve (Figure 31). This effect became apparent as participants in the study assigned to the Mechanix application consistently had statistically identical, or lower, scores compared to traditional homework on the first use of Mechanix, even when different homeworks were first. Consistent results were observed across each school, with a trend existing against the Mechanix condition on the first assignment. On the following assignment, participants at two of the three universities performed significantly better. This trend suggests that future deployments of the application could see more significant results if more problem sets are deployed as part of the Mechanix study.

One of the best presentations of this learning curve comes from the problem distribution at University 5. University 5 did not conduct two Mechanix problem interventions and instead distributes a single Truss problem set. The comparison of the Truss homework problem set at Univ. 5 compared to the other universities involved in the study exemplifies the learning curve, as it continues to show the trend in scores found at all other universities involved in the study on the FBD homework set. This presentation of the learning curve can be seen in Figure 31 below.
Figure 31: Data for first-time users of Mechanix tends to indicate a learning curve associated with the Mechanix homework system.

5.2 Influence of Mechanix on Student Achievement

While the general impact of the Mechanix application was not as pronounced as was expected, the performance of students that participated in the Mechanix application of the study seemed to indicate the application assisted in the quicker growth of student understanding of statics concepts. This impact can be seen in the trends and differences associated with the selected exam material at the different participating universities (Figure 10 and Figure 12). Participants assigned to the Mechanix condition showed better results immediately following the intervention. This difference disappears on later exams, which likely indicates a few things concerning the effect of the Mechanix application on student understanding. First, it seems to indicate that the application speeds up student understanding of the concepts, with the caveat of having a learning curve associated the interface usage. This revelation comes from the trend in which participants completing the first Mechanix assignment tend to have a significant number of questions about the
platform and seem to perform identically to participants who are assigned to the Control condition. A statistical difference at some schools between participants who completed the assignment using the Mechanix application indicates that once students are familiar with the application, the automated tutoring associated with the application enabled students to outperform their peers. Those students who were experienced using Mechanix outperformed those who were new to the application. While this difference cannot be attributed solely to the Mechanix application, the results of the other instruments seem to indicate that the Mechanix application does have an impact on immediate student understanding.

The interpretation that the Mechanix application has an immediate impact on student understanding is supported by students' performance on the collected exams throughout the semester, as well as the performance of students on the pre- and post-semester Concept Inventories. Following the intervention, participants assigned to the Mechanix conditions showed equal growth (University 2) or more significant growth (University 1) on concepts related to the material in Mechanix. This difference disappears later in the semester, which is evident in exam and CI scores collected after the Mechanix intervention has concluded. This erosion of impact suggests that high achieving students find the scaffolded instruction built into the application helpful as evidenced by higher exam scores immediately after they use Mechanix. Still participants assigned to the Control condition can catch up to the participants assigned to the Mechanix condition.

Starting with the exam results from University 1, student participants assigned to the Mechanix condition showed higher scores than students assigned to the Control condition (Figure 10). This difference disappeared on the final, suggesting that the
application may affect immediate student learning. The expectation is that students labeled as “high-achievement” will naturally self-correct, developing skills that may be missing following a lower exam score. As discussed in the results section, the Exam 2 Question 2 scores of participants at University 1 showed statistically significant results in favor of the Mechanix condition, suggesting that the Mechanix application has an impact on student learning.

University 2 across all exams shows that the Mechanix application does not harm students' understanding of the concepts (Figure 12). These results are consistent with the other instruments collected from University 2, indicating that the Mechanix application does not harm students' understanding of the material. Additionally, University 2 had high variation and small sample sizes, reducing the ability to detect differences in participant understanding. The lower population of the participating classes enabled this instructor to provide more detailed one-on-one instruction with participants, which is expected to have reduced the effects of the Mechanix intervention.

The inconsistent nature of the exam results collected at University 5 suggests that while the Mechanix application may have influenced student understanding, the controlling factor in student achievement seems to be closely tied to the time of day in which the polled section convened (Figure 13). The differences in the scores collected from University 5 show that no matter the intervention, the AM class performed worse than the PM class on Exam 1. This particular result seems to indicate that the time of day had a more significant impact on student performance than the perceived benefits of the Mechanix application. Once the discrepancy in the time of day was controlled, the difference in performance between the two conditions is consistent with the other
participating universities. Interestingly, after this trend was noticed, a conversation with the instructor at University 5 occurred concerning this difference. It was noted that only exam 1 showed this trend. For this reason, no major conclusions can be determined on the impact of the Mechanix application on student exam achievement at University 5, outside of the fact that it did not harm student performance.

These exam results suggest that the Mechanix application does increase student understanding of the complex concepts immediately following the implementation of the Mechanix intervention. When controlled for environmental conditions at the separate schools, the overall results from all universities showed that the Mechanix application did not harm student performance on exams collected as part of the study. These results were consistent at all universities where instruction and distribution of the exam were identical between semesters. After the Mechanix intervention has concluded, these differences disappear following the return to that course’s standard homework system. This suggests that while the Mechanix application may speed up student understanding of the material, student reactions to could cause them to match the knowledge of students that saw immediate learning benefits associated with Mechanix. Because of this pattern, it is theorized that an entire semester’s set of homework problems completed in the Mechanix application may have an increased impact on student understanding of the concepts. Additional exam data was collected and analyzed; however, the online testing that resulted in response to the COVID-19 pandemic made some of the exam data unusable compared to students who took in-person, closed-book exams.

On the subject of concept inventories, a few notable aspects were discovered. One of the essential instruments that students were assigned to complete was the FCI. It was
expected that students would have a solid grasp on the FCI concepts coming into the class, as the FCI covers topics traditionally introduced in physics classes. This was consistent with the results, as Participants at all Universities had almost identical scores between the pre- and post- distribution of the CI. This shows participants had an adequate understanding of Newtonian Physics concepts at the beginning, and that the inclusion of Mechanix in a participants instruction had no effect on their understanding of the material.

One of the main aspects that came out of the distribution of the CIs was that the high-level SCI may not target concepts taught by the polled classes involved in the study. There was also evidence that participants may have lost motivation on the lengthy CI. Additionally, the growth of a participating section was dependent on the curriculum of a class, with differences existing based on the exact curriculum and pre-existing knowledge of these student bodies. This difference was noted, and the order of the SCI was altered to reflect the expected difference in motivation as a student experience more challenging questions on the SCI after the Fall 2019 semester.

The reordered SCI survey was deployed to all schools following the Fall 2019 semester. The expectation was that students in conditions that previously showed lower effort in completing the survey would complete more of the survey, better reflecting the students’ overall understanding. Following the reordered distribution, a pronounced effect was discovered at University 5, which has a larger population of non-traditional part-time college enrollees. While these students still had lower performance than the participants at the other universities involved in the study (likely due to being a different major and curriculum), they performed significantly better than previous participants within the school. Comparing the Fall 2019 and Fall 2020 semesters shows that participants doubled
their scores from the earlier to the later semesters (Figure 18). This increase in score indicates that the reordered survey did impact student completion, meaning either more students had a greater understanding of the topics or more students had higher motivation in completing these surveys with the reordered questions. This difference is reflected in the other universities' scores through a trend in favor of the new order, but with no statistically significant difference.

5.3 Impacts of the Truss Design Problem

The results of the expanded problem distribution of the truss design problem suggest that the AI-augmented tutoring built into the application supports student achievement on an open-ended design problem. This claim is supported by the higher number of submissions, lower frequency of errors, and the prevalence of participants that use the Mechanix application to produce higher max load trusses than their peers assigned to the Control condition. These positive impacts came in fewer mistakes, higher max loads, and a higher number of submissions made by students at University 4.

Distribution of the Truss Design problem encountered some problems, starting during the Fall 2019 semester and continuing through to the Spring 2020 semester. It was assumed that the lack of relevant results from the pilot testing of the application was due to a few factors. The first is that the pilot problem statement deployed at University 1 was too simple, and the existence of the trivial solution was too obvious. The trivial solution showed that participants could solve the problem, although once they found a solution that worked, few attempted to exceed the minimum requirements. During the Spring 2020 semester, the distribution of the problem was directly impacted by the shift in education
due to COVID-19. The prevailing theory is that the change in education decreased the participating population, which is reflected in the results from University 1 show no significant difference. Some differences existed in the number of attempts made, although these results alone can not confirm that the Mechanix participants explored more of the solution space. It can be concluded that the introduction of the Mechanix application did not harm participants at University 1.

Full deployment of the application to University 4 received a good reception, resulting in rich data with a larger sample size. The Design problem results from University 4 showed a few aspects of the problem of note. The first being that the immediate formative feedback presented by the Mechanix application influenced participant understanding of the problem statement, coming from the difference in accurate responses between the two conditions. This significant difference could be due to a few notable aspects of the mass deployment of the design problem. First, either the problem was not deployed such that participants assigned to the Control condition did not fully understand the paper problem statement or the requirements for submission. The other consideration is that the scaffolded learning provided by the Mechanix application enabled participants assigned to the Mechanix condition to work through initial errors that may have interrupted solution attempts.

The corresponding max load results continue to support this theory, as the T-test related to the mean loads of the two conditions is significantly in favor of the Mechanix condition. The large sample size endorses the view that the lack of significance collected from the University 1 results can be attributed to the lack of participation. Through deployments of previous versions of the Mechanix application, it has been shown that
students at University 1 and University 4 produce comparable results [10-13, 80]. This previous comparability further suggests that the shortcomings of the University 1 data set can be partially attributed to the low participation rate due to the effects of COVID-19 on the study. Additionally, while the results of University 5 were inconclusive due to the lack of effort and possible excessive collaboration, the results continue to show that the Mechanix application enables participants to try many solutions in a shorter period. The results suggest that future deployments of the problem type at University 5 may see improvements using the low load problem statement, as no participants assigned to either condition produced a truss solution that met the design requirements.
CHAPTER 6. CONCLUSION

6.1 Addressing the Research Questions

As the focus of the experiment was on three specific research questions, addressing exactly how the results of the Mechanix distribution affected the hypotheses of the experiment is central to the evaluation of the academic impacts of the Mechanix application. For this reason, the three research questions have been included below:

**RQ1:** Does the usage of an AI-augmented tutoring system (Sketch Mechanix) improve student learning outcomes in Statics and Dynamics courses?

**RQ2:** Does the use of Mechanix affect the conceptual knowledge gains in a Statics course, as measured using established concept inventories?

**RQ3:** What are the impacts of using an intelligent tutoring software to solve an open-ended creative design problem?

6.1.1 Research Question One

RQ1 was developed to determine the overall impact of the Mechanix application, specifically if the usage of the application had a positive influence on student understanding of early mechanics concepts. This intention was to use the results from the homework interventions and the results from the collected exams to give concrete examples of the overall efficacy of the application’s academic impacts, which has proven to yield mixed results. The overall picture of student learning outcomes has been muddled due to the interruptions of academic studies in the last two years.
Homework results collected through the study have shown mixed results, with a learning curve associated with the application becoming apparent. Participating classes that deployed multiple Mechanix assignments experienced higher scores and an associated better learning outcome, as evident from the higher Mechanix scores collected from Univ. 1 and Univ. 4 during the second Mechanix intervention. Furthermore, the evidence of the learning curve can be directly seen by the difference in single intervention and double intervention schools, with the lower single intervention truss scores collected at University 5 contrasting with the scores at the other participating universities. Despite this learning curve, the difference in academic achievement due to the Mechanix intervention resulted did not harm student understanding of the material.

The evidence of the improved learning outcomes is supported by the exam results collected at the participating Universities. Participants assigned to the Mechanix condition at the minimum showed equal learning outcomes compared to members of the Control condition. The exam results immediately following the Mechanix intervention at University 1 seem to suggest that the application has immediate benefits in reinforcing critical concepts in engineering. It is theorized that the application provides an initial learning boost, due to the difference in homework and exam scores. Results from other participating universities show no harm done (University 2 and University 3) or show the influence of the class period's time on student learning outcomes (University 5).

As a response to RQ1, it can be said that the Mechanix application does seem to influence student understanding. However, the limited nature of the intervention reduces the overall impact of these results. It has been theorized that the effects of the learning curve may be minimized through the introduction of more homework assignments into the
application. At a minimum, a participant's usage of the Mechanix application did not harm their overall understanding, or learning outcomes, as shown through the graded material.

6.1.2 Research Question Two

The second research question focuses on evaluating the Mechanix application through comparisons of the performance of the Control and Mechanix conditions on standardized peer-evaluated concept inventories, specifically the FCI, SCI, and aDCI. The deployment of these CI’s across the participating universities as pre- and post- diagnostics resulted in comparative evaluations of the Mechanix application’s impact on student understanding and an assessment regarding the specific nature of these CI’s. The initial expectation was that participants of both conditions would see equal growth on the collected CI’s, with an additional theory that the Mechanix application may help on sub-categories of the SCI related to FBD concepts.

Collected CI results showed a range of impacts, with the distributed FCI’s no longer being used as a Co-variant in a more robust statistical analysis. The results showed that student performance on the abridged form of the FCI was identical between the Control and Mechanix conditions. These results support the initial theory that participants would have a solid conceptual understanding of force concepts entering the class, and their participation in the study would not harm their understanding. As the study matured, the results reflected this expectation. However, the differences in scores across the participating Universities reflected differences in understanding of Newtonian physics that may be due to an undergraduate’s major and, therefore, conceptual basis. Participants in
engineering adjacent majors performed worse than their peers enrolled in engineering-related majors, as evident by comparing results between University 1-4 and University 5.

The analysis of the Statics Concepts Inventory presented similar results, with the initial hypothesis dictating that participants would show equivalent results on the SCI in both assigned conditions. Furthermore, it was theorized that Mechanix participants would score higher on questions related to skills developed by the Mechanix application. This expectation was reflected in portions of the data, with the results suggesting that the Mechanix intervention had no effect on student performance. What did seem to hurt student performance was the difficult nature of the SCI. Results indicated that many participants, especially participants enrolled in non-engineering majors, lost motivation early into completing the SCI and produced lower scores than they could otherwise. The SCI was reordered from easiest perceived question subtype to hardest perceived question subtype, with results suggested that the loss in motivation lowering the score was indeed occurring. Additionally, the identical results between the two conditions at the participating universities indicate that either the SCI was not targeting skills that are improved by the Mechanix application or that the timing of the survey reflects the similarities found in the exam data.

As a short aside, the identical low scores of both conditions on the aDCI distribution at University 3 show the single intervention seems to have not targeted the concepts tested by this CI. The results indicate that both conditions showed no statistical difference and had equal conceptual gains. However, the low scores suggest that students either misunderstood the questions on the CI or the lower scores are evidence of low overall effort or knowledge. Regardless of the deficiencies in the collected data, it can be shown that a
student’s participation in the Mechanix condition did not harm their understanding of the concepts tested by the different CI’s. These results suggest that students who used Mechanix to complete homework showed equal conceptual understanding as those who used a traditional homework method. The results also indicate that the concepts covered by the different CI’s, especially the SCI, are not targeted by the concepts that are impacted by the Mechanix application. In short, the usage of the Mechanix application had little effect on conceptual understanding as measured by peer-reviewed Concept Inventories.

6.1.3 Research Question Three

The third and final research question was proposed to focus on the Mechanix application's capability to enable the distribution of a open-ended design problem. This question asked the participant to come up with a design meeting specific set requirements, which were altered depending on the specifications of the class that the problem is distributed to. The distribution of the design problem followed the condition assignment conducted for the other instruments used in the experiment, with Control participants completing the problem using pen-and-paper. The Mechanix application’s distribution of the design problem has been set up so that participants receive scaffolded feedback on mistakes made on entered solutions. It was theorized that this process would enable participants who use the Mechanix application to achieve better results on the corresponding design problem, as literature supports the expanded benefits of instant feedback in student achievement [3].
The design problem was distributed at three universities under two different distributions, the low load Pilot, and the high load distribution. The Pilot at University 1 used the low load problem statement and showed no improvements in student achievement that could be attributed to the Mechanix application. These results were inconsistent with expectations, as it was expected that the lower load requirements would encourage students to continue to explore the sample space once the first working solution had been achieved. This theory was discovered to be erroneous, and the impact of the Mechanix application on student achievement on the design problem was inconclusive. Following this Pilot, the design requirements were changed to reflect the conditions seen in Figure 4 of the results section. The pilot distribution did confirm the stability of the Mechanix application’s version of the design problem, allowing for future distribution.

After the pilot problem failed to produce benefits from Mechanix, all following distributions followed the high load problem statement. The design problem was extended to all participating Universities, with three Universities deploying the high load problem statement. The high load problem results were more consistent with the original Research Question 3 hypothesis. Specifically, the hypothesis that the instant feedback and scaffolding provided by Mechanix allows participants to create final truss designs with a higher max load. Results from University 4 suggest that the application provided students with a significant amounts of support, both in creating better final solutions and reducing the number of errors made (by providing instant feedback). The difference in number of solutions between the two experimental conditions (Control and Mechanix) at University 1 suggests that the application’s ease of access permits participants to explore more of the solution space, as Mechanix participants produced a larger number of attempts than their
peers in the Control condition. The results from both University 1 and University 2 confirm parts of the original hypothesis, specifically that participants would be able to explore more of the design space and result in higher achievement when compared to their peers who complete the problem using pen-and-paper.

6.2 Conclusions on Mechanix Homework

While it cannot be concluded that the Mechanix application is outright better than traditional (online or pen-and-paper) homework methods, it can be said that the application may have an immediate impact on student learning of these subjects. Through the continued deployment of the application, it has been discovered that the continued usage of the application affects student understanding, although with an initial learning curve. Once a student overcomes this learning curve, the scaffolded approach to instruction developed through automated personalized feedback can improve student understanding of the material faster than traditional homework methods. This initial learning curve associated with the application seems to impede initial student understanding.

Future steps in evaluating the impacts of Sketch Mechanix are expected to follow the implications of this learning curve, as an entire classroom set of problems is being created for Mechanix. This problem development is intended to grow the Mechanix application's reach in influencing student understanding of statics concepts, as it will enable faculty to use Mechanix throughout the entire semester. Future distributions of the Mechanix application are expected to include more problem sets, which is expected to have a more pronounced influence on students than distributions with a limited number of questions.
6.3 Impacts of the Truss Design Problem

While the initial results from the full deployment of the Mechanix truss design problem are promising, future deployment of the application is expected to produce richer data. The existing results from participating universities show that the application has at the very least had an equal effect on student submissions as a pen-and-paper submission method. Of important note was the significant difference in the total number of submissions between the Control and Mechanix condition, favoring the Mechanix submission rate. It is believed that this difference can be attributed to the Mechanix application providing support for the student submissions, enabling students to explore the solution space to a greater degree and providing immediate feedback. This difference is intended to be tested with the future deployment of the design problem. These deployments will put additional emphasis on Control participants submitting more than only their final results. The intention is to show that the overall difference in submissions per student can be attributed to exploring the solution space instead of simply a lack of submitted attempts made by members of the Control condition.

It was also found at University 4 that student participants that used the Mechanix application significantly outperformed their peers in the Control condition, producing final solutions that could hold a larger max load. The leading theory concerning the Mechanix application allowing students to explore more of the design space is supported by the participants at University 4 producing truss systems with higher max loads. Due to these current results, it is expected that further deployments of the Mechanix Truss Design problem at University 1 would yield similar results, as an increased population would produce a larger participant sample size. It has also been proposed that the low load
problem statement might show similar results at University 5, with the Spring 2021 results reflecting a lack of motivation in completing such a complex problem.

Overall, the deployment of the Mechanix Truss Design problem shows that the application has a significant impact on student understanding of the material, with tutoring built into the application providing students with the opportunity to come up with better results than their peers in the Control condition. The tutoring in the Mechanix application has been shown to provide immediate, meaningful feedback in large populations, as evident from the solutions made by Mechanix participants at University 4. The overall rate of correctly calculated answers from the Mechanix condition at University 4, as opposed to the number of erroneous submissions of the corresponding Control condition, supports one theory that the tutoring of the application assists in understanding the design problem. Furthermore, the overall rate of positive submissions at all universities is consistent, meaning large amounts of the Mechanix population at all participating schools produce working truss solutions. These findings suggest that the truss design problem built into the Mechanix application may act as a form of active learning that could be distributed to early Mechanix participants and is a possible solution to the lack of these hands-on assignments in early engineering course work.
6.4 Limitations

While the study has managed to produce analyzable results during each semester, limitations became apparent as the study progressed. One of these limitations is based on the timing of the distribution. Another limitation of the distribution methods is assigning an entire section to one condition instead of splitting sections into both conditions, especially when the different distributions occurred across multiple semesters. Limitations due to differences in the curriculum from semester to semester, differences in exams and instruction from section to section, and uncontrollable natural impacts from semester to semester have directly impacted the efficacy of portions of the data sets. These limitations are not always controllable, but the evidence that split sections produce more comparable results than single condition sections can and will be considered in future studies on the academic impacts of the Mechanix application.

Another limitation comes in the form of interruptions to classroom instruction due to COVID-19. As most classrooms acted in a distance learning fashion for most of the study, major differences in participation rates occurred at the polled Universities. Additional impacts due to other climate-related incidents impacted data collection, as freak blizzards and wildfires occurred during multiple semesters of data collection. To compound with this, unintended shifts in the class instructor during the two main semesters of data collection may have also influenced the rate of participation in the study. Due to the influence of both shifts in instructors as well as COVID-19, many of the student participants who initially signed up for the study eventually declined to participate, decreasing the usable study results.
A significant limitation discovered in the truss design problem distribution was the existence of a trivial solution to the problem statement. This trivial solution became a serious issue during the pilot distribution of the problem, as it decreased the richness of the participants' answers. During the full distributions of the problem, the challenge of the problem statement was increased by altering the criteria, invalidating the original low load problem’s trivial solution. Despite this, by simply mirroring the trivial solution about the X-axis, the truss can handle close to double the load of the original trivial solution. Participants of either condition did not find this new trivial solution at the same rate as in the pilot study, suggesting that the new problem statement encouraged students to explore more of the solution space than the low load problem statement. After internal discussion, the leading theory behind why this difference in cognition occurs is how instructors often teach using a traditional truss design and avoid describing deck truss geometry. Students internalize this method of truss creation and create solutions that match their pre-existing expectations. For future distributions of the problem, the high load problem statement is expected to produce rich results at high achieving universities (Universities 1 and 2). In contrast, the low load problem statement is expected to produce better results among non-engineering majors (University 3). However, alterations to the problem statement at participating universities may limit the impact of any trivial solution or difficulties experienced by participating students.
6.5 Future Work and Reflections

The distribution of the Mechanix application that is the focus of this thesis is not intended to be the final distribution of the application. A tremendous amount of effort has been put into ensuring that the application can see expanded deployment in classroom settings after the current research. These efforts to build a more robust Mechanix application have proven fruitful, as parts of the application have been deployed in classes outside of the original purview of the grant. This has been conducted to get the application into more students’ hands and allow more students to experience the benefits of the application. Plans have been made to continue to deploy the application in future sections of the class past the departure of the current research assistants.

My experience with the application has been insightful, with conversations with participants and my own usage reflecting an overall positive opinion on the impact and use of the Mechanix application’s capabilities. The level of care and attention that has gone into creating the application shows me that while engineering education may change over time, the instructor's goal is always to help students become the next generation of engineers. The tools that we can develop now can help instructors and students of the future, and it is my belief the problem types that can be distributed using the improved architecture can meet the requirements of the future.
APPENDIX A. MATERIAL DISTRIBUTED TO PARTICIPANTS

A.1 Consent Form

As part of the study, students were informed of their opportunity to participate through a consent form, which was distributed during the first class period. While no script was used for the study, the consent forms were consistent across the many distributions of the study. Below is the form distributed as a webform through the Mechanix application. Some personal details have been removed to preserve anonymity.

(webform when first logging into the software)

University 4 HUMAN SUBJECTS PROTECTION PROGRAM INFORMATION SHEET

Project Title: Enhancing Visualization Skills and Conceptual Understanding Using a Drawing Recognition Tutoring System for Engineering Students

You are invited to take part in a research study being conducted by Dr. #######, Dr. ######, and Dr. ######. The research is funded by the National Science Foundation Award # #####. The information in this form is provided to help you decide whether or not to take part.

Why Is This Study Being Done?

The purpose of this study is to understand what types of approaches enhance student’s learning of statics and mechanics of materials. The Mechanix software is designed to
provide immediate and personalized feedback on your hand-drawn homework solutions.

This research project is trying to see how effective the software is in aiding understanding for differing groups of students based on preparation and other demographic information.

The researchers also want to obtain feedback as to how to make it better.

**Why am I being asked to be in this study?**

You are being asked to be in this study because you are enrolled in a course that is testing out educational software the provides immediate personalized feedback to student’s hand-drawn solutions.

**How many people will be asked to be in this study?**

2,000 people (participants) will be invited to participate in this study locally. Overall, a total of 10,000 people will be invited across multiple educational facilities.

**What will I be Asked to Do?**

To participate in this study, you will perform your normal classroom activities. Your teacher will introduce the concepts and assign homework. A portion of the class (up to 100%) will be randomly chosen to use the Mechanix software that provides immediate and personalized feedback on your solutions. All participants will use Mechanix to complete a pre- and post- test and the optional demographics questionnaire. The teacher will introduce Mechanix and how to use it to submit your homework. After the study is complete, even though students who were not chosen to be in the test group will be given the option to use Mechanix to study for their final exam.
Note that you still need to do your homework whether or not you participate in our research study. That means you may still have to use Mechanix to submit your homework, it will just not be included in our study.

For our research, we ask your permission to use the data from your software data to help us determine whether Mechanix is an effective tool for studying. Through Mechanix, you will also be requested to supply optional demographic information, such as your age, grade, major, and prior experience. As part of your class requirements, Mechanix data will also include pre- and post- tests of your knowledge of the subjects as well as your scores and your solutions to the assignments. You may also opt to provide verbal comments about the software if you so with in a focus group.

If you decide to participate in a focus group, it will be recorded with an anonymous voice recorder so that your anonymity will be preserved.

**What are the possible benefits of this study?**

The study will provide insights as to how to create better educational software that can improve student learning.

**Do I have to participate?**

No. Your participation is voluntary. You may decide not to participate or to withdraw at any time without any effect on your class grade or any other relations.
Again, note that you still need to do all course-required homework assigned, including the pre- and posttest whether or not you participate in our research study. That means you may still have to use Mechanix to submit your homework, it will just not be included in our study.

Who will know about my participation in this research study?

Your course instructor will not know whether you have chosen to participate in the study or not until after the semester is over through the use of a computer time-lock.

Will I be compensated?

No.

Will Information From This Study Be Kept Private?

The records of this study will be kept private. Research records will not contain any identifying information. No identifiers linking you to this study will be included in any sort of report that might be published. All research information will be stored in a password-protected database behind a secure firewall at Texas A&M University.

People who have access to the collected information include the Principal Investigator and research study personnel. Representatives of regulatory agencies such as the Office of Human Research Protections (OHRP) and entities such as the University 4 Human Research Protection Program may access the records to make sure the study is being run correctly and that information is collected properly. The agency that funds this study (National Science Foundation) and the institution(s) where study procedures are being
performed *University 4* may also see your anonymous information. Any information that is sent to them will be coded so that they cannot tell who you are. If there are any reports about this study, your name will not be in them. Information about you and related to this study will be kept confidential to the extent permitted or required by law.

**Who may I Contact for More Information?**

You may contact one of the Principal Investigators to tell them about a concern or complaint about the research.

For questions about your rights as a research participant; to provide input regarding the research, or if you have questions, complaints, or concerns about the research, you may call the *University 4* Human Research Protection Program office by phone at ####, toll free at #####, or by email at email@email.edu.

**What if I Change My Mind About Participating?**

This research is voluntary and you have the choice whether or not to be in this research study. You may decide to not begin or to stop participating at any time. If you choose not to be in this study or stop being in the study, there will be no effect on your class grades, student status, relationship with *University 4*, etc. To change your mind, click Settings/Consent in the menu in the top right.

**STATEMENT OF CONSENT**

The procedures, risks, and benefits of this study have been told to me and I agree to be in this study. My questions have been answered. I may ask more questions whenever I want.
I do not give up any of my legal rights by signing this form. You are not being asked to sign this consent for since your signature would be the only record of your participation. Rather, please click the appropriate button below to specify your intention to participate. If you would like to participate in this study click here: YES, and I am 18 or over If you would not like to participate, click here: NO

A.2 Tutorial Video Email

As part of the study, participants were informed of a series of tutorial videos that were created to provide participants with meaningful understanding of how the application behaves. The email that was sent out to the students is included below, with some information removed to remain anonymous.

Good afternoon,

As discussed in class your next homework assignment “Problem Set 5” is on Sketch Mechanix, not WileyPlus. On WileyPlus you will see that the homework assignment was due 5 minutes after the assignment was assigned, as you are assigned the problem set on Sketch Mechanix. On Thursday I showed all of you the rooms in MRDC that you are free to use for the completion of thes problem set, you are not required to use these rooms to complete the problem set. The code to get into this room is listed below:

MRDC ####

The assignment on sketch mechanix is due Thursday at 10 PM. Should you need additional help in understanding how to use Sketch Mechanix, I have attached some links to tutorial videos on Sketch Mechanix.

General Mechanix Video: https://www.youtube.com/watch?v=kaTlhcvGSNs&feature=youtu.be
Truss Video: https://www.youtube.com/watch?v=POpQYPyhyhNI&list=PLImvSFjmqX6rr7OaAFSSdqc0mSve0ur y2&index=2
Mechanix Playlist: https://www.youtube.com/watch?v=tPs9E76TOYw&list=PLImvSFjmqX6rr7OaAFSSdqc0mSve0ur y2&index=1

If you have any questions please do not hesitate to reach out to me.
A.3 Mechanix Instructions

Step 1: Read the problem statement and begin drawing a Free Body Diagram (FBD).
- Mechanix will change the color of your sketch to orange and add nodes (lettered key points) when you draw an appropriate sketch.
- Mechanix may provide a light background image to assist sketching.
- If Mechanix does not recognize your sketch, consider including different parts of the system in your sketch.
- Stray marks may prevent Mechanix from recognizing your sketch, even after you erase them. You may need to delete the entire sketch by clicking the trash can icon.

Step 2: Draw all forces acting on the body and label the forces in the text boxes.
- Your sketch must be orange before drawing forces.
- Correctly drawn forces will change color to purple and allow you to label them.
- Forces may be drawn with a single stroke or multiple strokes.
- Arrowheads that are symmetric and sized proportionally are more likely to be recognized.
- Forces must begin or end at a node.
- Forces must have different labels.
- Forces at an angle are decomposed into x and y components. Use these components when entering equilibrium equations (do not use sin or cos).

Step 3: Submit the equations of equilibrium and unknown values.
- You may use the sketching area for scratch work after the sketch is recognized in orange.
- Make sure your signs (+/-) agree with the directions of forces in your FBD.
- Be sure to include units.
- Mechanix will reveal correct and incorrect answers after clicking the “Submit” button (see figure on right).
- You may submit answers up to three times per problem.

General Information
- The “Save” button will save your progress but will not submit for grading.
- If you have questions or need technical help, please contact [grad student] at [email address].
- Send software bugs to mechanixbugs@googlegroups.com

Updated: 9/6/2019
APPENDIX B. EXAMPLE DISTRIBUTED PROBLEM SETS

B.1 Distributed Homework Problems

The following appendix contains a sample set of the problems distributed to students. Out of the traditional homework, only the formatting for the paper distribution is included. All other forms of the traditional homework distribution are simple format changes to match the online system that is used at the participating universities.

The WileyPlus homework problems were used to test the Mechanix application. Some of the traditional homework problems distributed as part of the study came from the Statics course built into WileyPlus and the associated textbook. The question numbers have been included for the associated textbook used during the experiment, specifically the Eighth edition of the Statics textbook printed for Georgia Tech as well as problems included in the MasteringEngineering system by Pearson.

B.1.1 Homework Distribution, FBD Problem Set

WileyPlus eighth edition Questions:

- Chapter 3, Problem 3/010
- Chapter 3, Problem 3/018
- Chapter 3, Problem 3/030
- Chapter 3, Problem 3/034
- Chapter 3, Problem 3/044
MasteringEngineering Questions:

- Chapter 5, Problem 13
- Chapter 5, Problem 19
- Chapter 5, Problem 26
- Chapter 5, Problem 30
- Chapter 5, Problem 44
- Chapter 5, Problem 45
- Chapter 5, Problem 85
- Chapter 5, Problem 92

B.1.2 Homework Distribution, Truss Problem Set

WileyPlus eighth edition Questions:

- Chapter 4, Problem 4/01
- Chapter 4, Problem 4/05
- Chapter 4, Problem 4/06
- Chapter 4, Problem 4/09
- Chapter 4, Problem 4/010
- Chapter 4, Problem 4/014

MasteringEngineering Questions:

- Chapter 6, Problem 4
- Chapter 6, Problem 5
B.2 Truss Design Paper Problem Statement

In this assignment, you will work on your own to design a truss to meet a set of requirements. Specifically, you will be designing a small bridge for hamsters to make their cage a more exciting and stimulating play place. While your design will need to support a minimum load in the configuration described below, the student that designs the truss that can hold the maximum load will receive a small extra credit bonus on the assignment.

Truss Design Problem Requirements

1. Your design must be a truss supported on one corner by a pin and another corner by a roller
2. The truss should span between 10.9 and 13.4 cm between the supports
3. The truss must hold a minimum load of 13N. This load is divided evenly across all joints on the top of the structure.
4. Each truss member can hold a maximum of 4.5N in compression and tension.
5. Each node across the top of the truss shares the loading equally. For example, if there are 3 nodes across the top of your truss, each node will have an applied load of P/3 downwards.
6. All distances you input are in cm. All angles are in degrees.

What you will do

1. On your own, brainstorm design ideas on paper. Analyze these designs using the methods you have learned in class.
2. For your designs, determine the max value of P before the truss fails.
3. Continue to improve your design, submitting each major iteration into the program, until you are satisfied with your final submission.
4. In canvas, submit the following:
   a. Your by-hand analysis for your final truss
   b. The finalized Truss FBD clearly marked with dimensions and loads
   c. An appendix that includes your previous iterations and significant analyses done outside of the design tool.
APPENDIX C. TRUSS DESIGN CODEBOOK

The following appendix details the codebook used while assigning feedback to participants who submitted an answer to the truss design problem under the Control condition. This codebook was used for the Inter-rater reliability.

<table>
<thead>
<tr>
<th>Feedback Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solved Truss but did not meet requirements</td>
</tr>
<tr>
<td>Invalid Truss Supports</td>
</tr>
<tr>
<td>Invalid Truss Geometry</td>
</tr>
<tr>
<td>Unstable Truss</td>
</tr>
<tr>
<td>Truss could hold larger load</td>
</tr>
<tr>
<td>Incorrect calculated load</td>
</tr>
<tr>
<td>Did Not include work</td>
</tr>
<tr>
<td>Incorrect application of load</td>
</tr>
<tr>
<td>Incorrect Assumptions Made</td>
</tr>
</tbody>
</table>

Submissions are coded as independent submissions if they created distinct geometry and made attempts at solving the truss. If the student simply provided valid truss geometry but did not attempt the solution, this does not count towards the total number of submitted trusses. The only exception being if they did not submit work. That is on a case by case basis and subject to coder discretion.

Remember to check the Truss geometry prior to looking through work. Incorrect Truss Geometry is assumed to be the initial failure point for the solution attempt.

Submissions Coded on the Hierarchy of:

Solved Truss but did not meet requirements

Feedback is usually assigned based on the student solving the truss correctly, but not creating a truss that does not meet the minimum load requirements. Additionally, this can be assigned if the truss is solved correctly, but the student did not meet the length requirements.

Invalid Truss Supports

Usually this occurs when the two supports are not on the same elevation, as they are asked to create a planar truss that bridges a flat gap. This usually doesn’t occur and may be a limitation of the current problem type.

Invalid Truss Geometry/ Overconstrained
This usually happens when a student attempts to create planar truss geometry that does not fit the criteria. Specifically this occurs when a student has two members cross when it is not a joint. Other situations are were a student creates a statically indeterminate truss.

**Unstable Truss/ Underconstrained**
Assigned this feedback when a truss does not meet the x-y stability requirement. The problem statement should have been modified to ensure that this is not allowed, but students still attempt to submit this.

**Truss could hold larger load**

This is usually given, but not really considered a downside. The student met the minimum requirements, but just doesn’t solve the maximum applicable load. Usually counted as getting the problem correct for analysis purposes, if their solution meets the minimum requirement for the system.
Incorrect calculated load

This is an identical truss as before, with an over calculated load. This is given to a student that either under or overcalculated the load based on a math mistake.

**Did Not include work**

Self Explanatory. Student submitted a proposed truss, but did no work to solve. This usually results in the inclusion of the truss submissions (that appear fleshed out), but does not count towards a correct submission.

Incorrect application of load
Usually comes in the above form, where students do not apply the load at a joint, instead they treat a truss member as a beam, and apply a load at an incorrect location. Additionally, students do not apply their loads to the correct number of locations, or decided to change the magnitude across the truss members.

**Conceptual Error Made**

Assigned if the student makes a major conceptual error in calculating the truss, or misunderstands how the load is supposed to be applied to the system.
REFERENCES


J. L. Davis and T. McDonald, "Online Homework: Does it Help or Hurt in the Long Run?," 2014.


