A Study on the Impact of a Statics Sketch-Based Tutoring System Through a Truss Design Problem

Josh Hurt
George W. *Woodruff School of Mechanical Engineering*
Georgia Institute of Technology
Atlanta, USA
jhurt30@gatech.edu

Matthew Runyon
Department of Computer Science & Engineering
Texas A&M University
College Station, USA
mattrunyon@tamu.edu

Tracy Anne Hammond
Department of Computer Science & Engineering
Texas A&M University
College Station, USA
hammond@cse.tamu.edu

Julie S Linsey
George W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, USA
julie.linsey@me.gatech.edu

***Abstract*— Providing opportunities for students to exercise their creative skills in large, entry engineering classes challenges most faculty. This paper presents a study of a large statics class provided with a homework problem that asks them to design a truss structure. Automatic grading was done by Mechanix, an AI tutor-based software package that can automatically recognize a free-body diagram or a planar, 2d, statically determinate truss structure. The paper presents a study done in two different semesters, comparing the students using Mechanix to a control (problem on paper). To ease grading, the control group's trusses were analyzed by Mechanix after submission. No mean homework grade differences were observed, but students in the Mechanix group produced trusses that could withstand higher loads. This is despite the fact the only guidance or feedback Mechanix provides was if the students' calculated max load was correct, and if it was not, which member failed. This study occurred in Fall 2019 and Spring 2020. Students also submitted more attempts in Mechanix than the control. It may be students in the control group only submitted correct answers despite being asked to submit all attempts. Future work will provide more incentive for students to submit all attempts on paper. Mechanix automatically records all attempts. During high stress (Covid-19), more students in the Mechanix group submitted the assignment indicating that students may find this system less mentally taxing to use, less stressful, or something else led to this difference. It will be explored with focus groups in the future. AI tools have the potential to provide automatic grading for open-ended, creativity required, design problems, and to engage students more, allowing universities to develop more innovative engineers while also deepening their knowledge.**

***Keywords*—Free Body Diagram, Statics, Sketch Recognition, Homework**

I. INTRODUCTION

To deepen students' knowledge and support their skills in innovative with that knowledge, students need more opportunities to use their engineering science creativity, yet this is often very challenging in large, entry-level engineering courses. Creativity sits at the highest levels of Bloom's taxonomy, indicating that knowledge is deep when students can be creative with the knowledge [1]. The Engineer of 2020 calls on engineering educators to build students' innovation skills [2, 3]. The graduating engineering class of 2020 already sit in our classrooms, and much more needs to be done to develop the highly innovative, creative engineers needed to solve global warming, minimize environmental impacts, address health crises, overcome world hunger and many other challenges faced by the world today. Students need open-ended design problems throughout their engineering curriculum and not just in the very limited design classes. Engineering educators desire to meet this goal, but providing these learning opportunities in lower-level, typically large classes is often very difficult due to class size. Well-developed, AI systems have the potential to allow open-ended design problems to be automatically graded, allowing instructors to assign these problem types even in large classes. Mechanix has been developed to enable students to design 2d, planar, statically determinate trusses, and then to provide automatic feedback and grading.

II. MECHANIX

Mechanix is an online Free Body Diagram (FBD) tutoring application, which allows students to get more personalized instant feedback on drawn FBDs, as well as the entered answers (sketchmechanix.com, email authors for access). It is currently a non-commercial software package. The software was initially designed to provide feedback for problems related to trusses but has since been extended to more problem types, ranging from simple Statics situations to more complex Dynamic Equilibrium problems. The application provides more feedback than traditional online homework systems while giving far quicker

National Science Foundation Grants 1725423 and 1726306.

feedback than pen-and-paper homework submissions. The newest addition to the software is a Truss design tool, which provides an interface to draw a truss system to meet specific requirements set by the instructor. Figure 1 details the user interface used by Mechanix, in which students are prompted with a specific problem, and provided with a field to draw the problem free body diagram (FBD) associated with the problem, while also providing answers needed to solve the given problem.

Most online homework systems do not require students to provide the FBDs used to solve the problem. The result is a situation where students do not always practice drawing FBDs on easier problems, which may cause conceptual issues when the student approaches the more complex problems. Mechanix requires the student to draw by hand the FBD associated with a problem, before the student provides the answers to the problem. Figure 2 gives an example of the drawing surface with a user-submitted Truss system. Students use a tablet-based computer and sketch like they naturally would on a piece of paper. The Truss shown below fails to meet the problem requirements but provides an example of the range of freedom that students are allowed in solving the problem.

III. Background

Too often, faculty observe students failing to create FBDs, a critical external representation for reducing a real-world entity to a model that can be analyzed. FBDs externalize a student's reasoning about a system. Through externalizing it, both the faculty member and the student can review and reflect on the model created. With the necessary large class sizes in engineering, faculty often implement online learning systems that automatically grade assignments. These systems are highly effective for providing large numbers of practice problems for the students and automatic grading. Similar systems to Mechanix for statics do exist. These packages include WinTruss [4], McGraw Hill Connect Engineering, Bridge Architect[5], Free-Body Diagram Assistant [6], and TrussMe! (phone app) [7]. Most are palette-based systems where users can pick pieces and use a mouse to drag them on the workspace to build their solution. Newton's Pen allows users to draw the FDB but constrains the user to draw free-body diagram components in a very specific order [8]. Unfortunately, none can provide feedback on natural, free-hand sketched FBDs or recognize a sketched truss and evaluate the open-ended design of it.

Mechanix, including the Truss Design Mode, provides immediate feedback to the students. Immediate feedback plays a crucial role in improving learning [9]. Feedback helps correct incorrect concepts and identifies mistakes learners are making. In a systematic review of the literature, formative feedback, (in contrast to summative feedback) has emerged as the most valuable type of feedback [10] and Mechanix provides this. A meta-analysis concluded that immediate feedback provides greater benefits as opposed to the delayed feedback that often occurs [11], and these results have also been found with undergraduates [12].

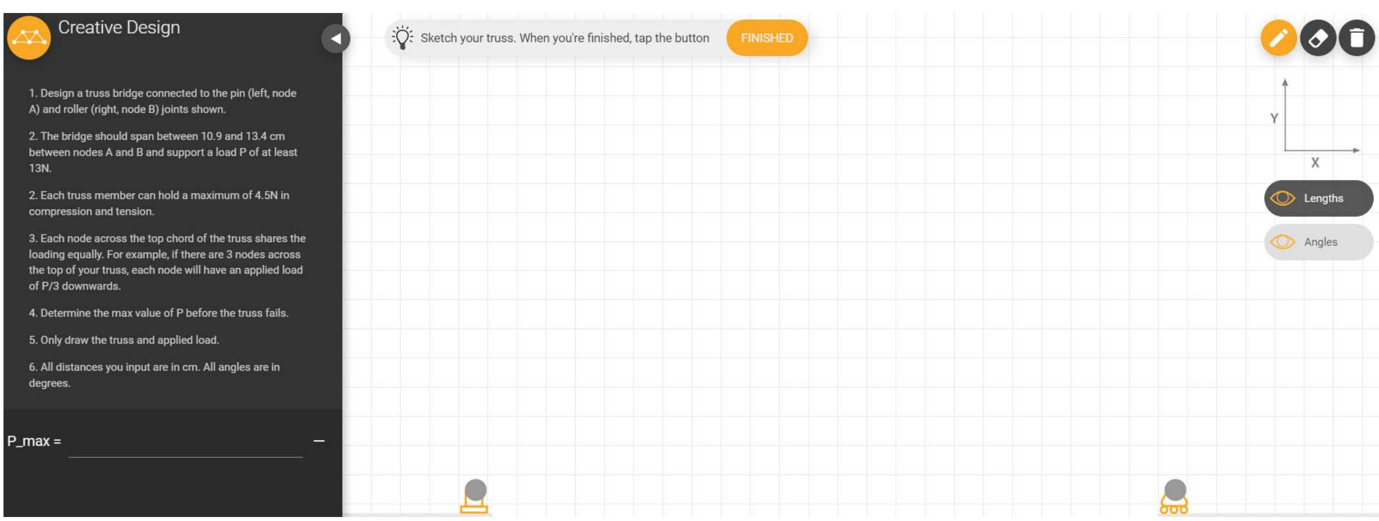

Figure 1: An example of the Mechanix User Interface, in which nothing has been drawn. This is what greets the user, providing the problem prompt, as well as drawing tools in the top right corner, with a large area to draw the truss system related to the problem.

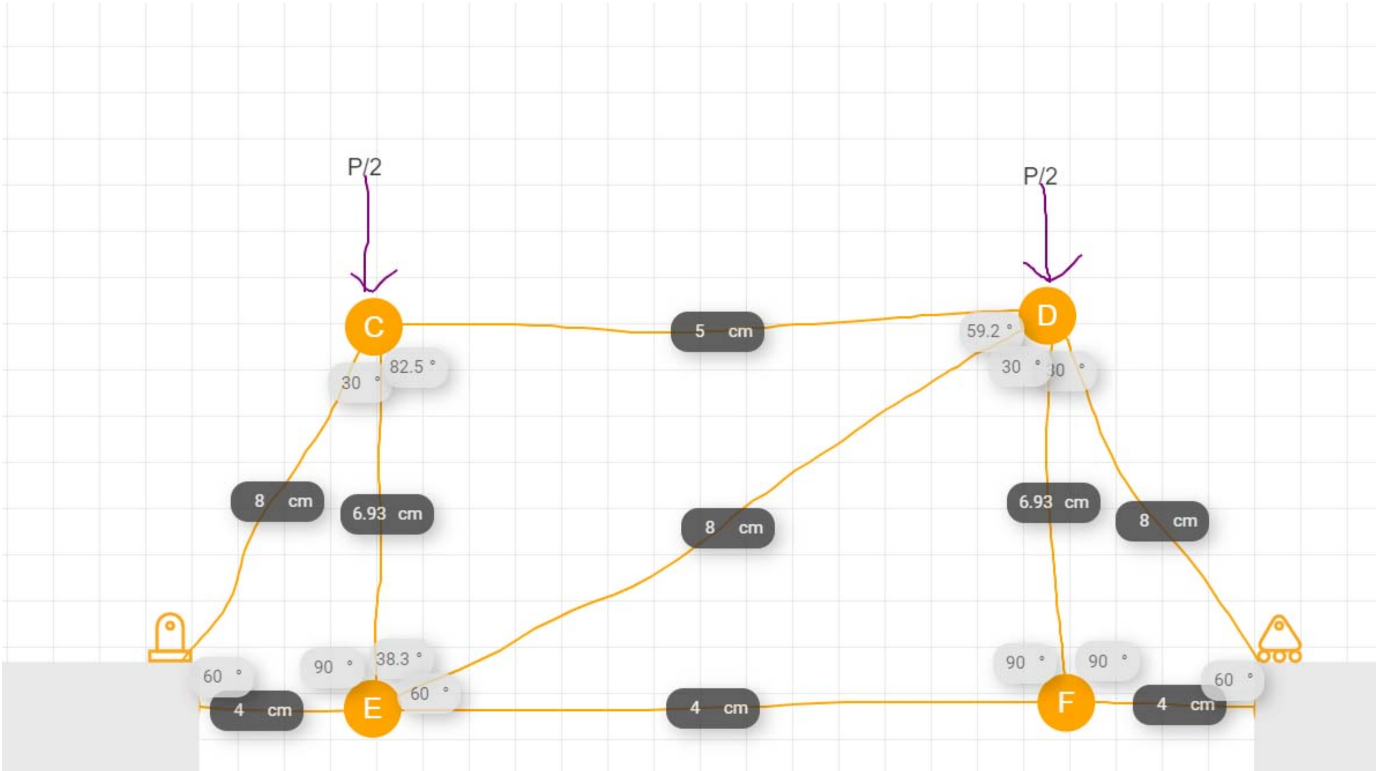

Figure 2: Example of Mechanix User Interface, with a drawn Truss system. This example comes directly from the Truss Design Mode, in which students are given freedom to develop their own solution to an open-ended question, in which they have to meet certain specifications set by the instructor while attempting to maximize the load that the truss can take.

IV. Methods

Students in the same section were randomly divided into experimental and control groups, where the control condition completed homework assignments using WileyPlus and the experimental group completed specific homework assignments using Mechanix. The use of Mechanix throughout the semester is a part of a larger multi-university study and this paper focuses on the Truss Design Mode only. The homework assignments were mandatory assignments and agreeing to participate in the study would net students extra credit towards their homework grade. Only data from those who agreed to take part in the are included in this paper. This design problem tasked students to create a truss system that would meet minimum specifications for the overall truss length, allowable internal member forces, and length of truss members. The control group completed the assignment on paper, while the experimental group used Mechanix. The submissions were graded based on if the submitted truss system met the criteria of the assigned problem and if they correctly calculated the proposed maximum load. Students were told that whoever submitted a truss that could take the highest total load would be awarded extra credit points towards their overall homework grade.

The control group was asked to submit the iterations of their work that lead to the final solution, as well as the maximum load that their final submission could take. The experimental group submitted their assignment on Mechanix, which allowed a student to draw an open-ended truss system and input their calculated max load into the system. Mechanix collects all submission and would then give students a range of feedback depending on the submitted answer. This ranged from warning the student that the drawn truss was incomplete or incorrect, to providing insight into which members of the truss would fail first. Both groups were directed to take as many attempts at the problem as they wanted, to prime students to explore the design space of the problem. This meant that students were allowed to submit as many solutions as they wanted both on paper, as well as through the Mechanix.

V. Results

Mechanix has been extensively tested in the classroom environment for some time now, with past results showing that the system performed as well as traditional homework methods [13-17] . However, the new Truss Design Mode has only been tested twice so far in classroom settings, once in the Fall 2019 semester, and then again during the Spring 2020 semester. The first usage of the Truss Design Mode in a classroom setting was a test run, with the intention of gathering user feedback for the problem and software.

The initial testing of the Truss Design problem occurred in Fall 2019 at the Georgia Institute of Technology in an introductory Statics class. The simplified problem had minimum requirements, which should be obtainable with some of the most obvious truss designs. The student-designed truss had to span between 10.9 and 13.4 cm, was required to carry a minimum load of 7.5 N distributed evenly across the top nodes of truss, members be 4.2 cm long, and the internal member force to not exceed 4.5 N. Students were allowed to add one additional member length of their choosing.

In initial testing, this simplified problem was found to have a trivial solution, where only equilateral triangles could be used to produce a truss system that would meet the requirements, however, it was decided this initial problem description would be used to test the software. As a result, the students generally found the trivial solution, and little iteration occurred in the development of the problem submissions. This resulted in the mean grade for the assignment between the experimental and control groups being almost identical, which is shown in Figure 3 below, and zoomed in Figure 4. A two-sample t-test assuming unequal variances, demonstrates the grades for the two groups are not statistically different (t-value=0.132, dof=35 P=0.895).

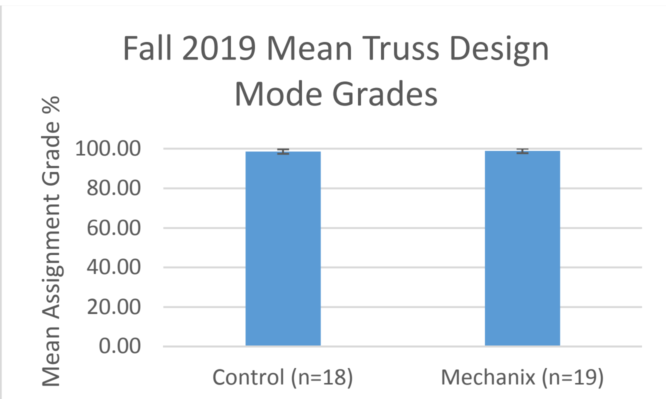

Figure 3: F19 Georgia Tech Mean Truss Design Mode Grade by group. The simplified Truss Design Mode resulted in most students understanding the material, however it showed that the problem in its current form was not challenging enough.

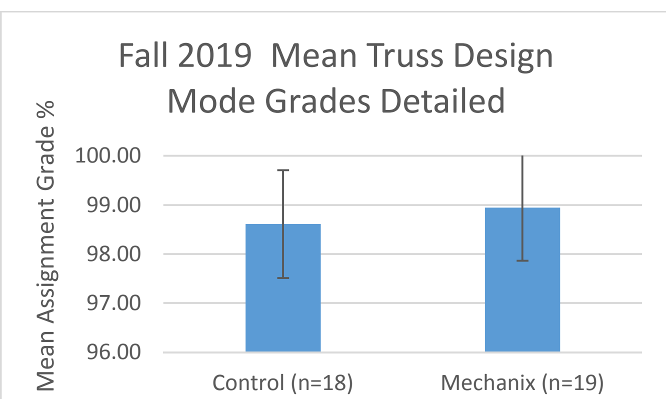

Figure 4: F19 Mean Truss Design Grade by group. The grade results of the simplified Truss Design Mode zoomed in to show no distinction between the grades.

Another item that was recorded for the Fall 2019 Truss Design Mode was the maximum load that the student submitted trusses could handle. This was reported in Newtons and was one of the required parts of a student's submission for grading purposes. Mechanix checked this answer automatically. As noted before, students were incentivized to achieve a higher permissible load through giving extra credit to the student that achieved the highest possible load . Despite this, most students stopped attempting the problem after they found a truss that would meet the requirements. Out of the students that participated (n=37) in the study, 7 students between both of the groups attempted the extra credit aspect of this assignment and went beyond the trivial solution. The experimental group had a mean maximum load of 7.98 N, while the control group had an mean maximum load of 8.59 N. A t-test assuming unequal variance was conducted to test if there was any statistical difference, showing that on the simplified problem statement there was no statistical difference between the performance of the two group conditions (t-value=1.512, dof=26 P=0.142). This trend is exemplified in Figure 5 and further exemplified in the zoomed-in Figure 6.

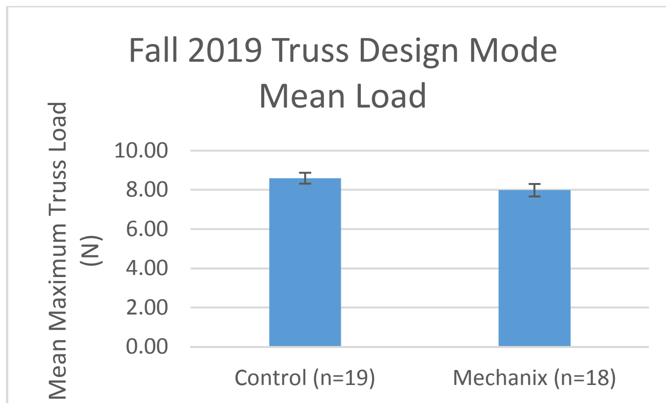

Figure 5: Fall 2019 Mean Truss Design Mode Truss Load. The mean load applied to the student submitted truss, broken down by the experimental and control groups.

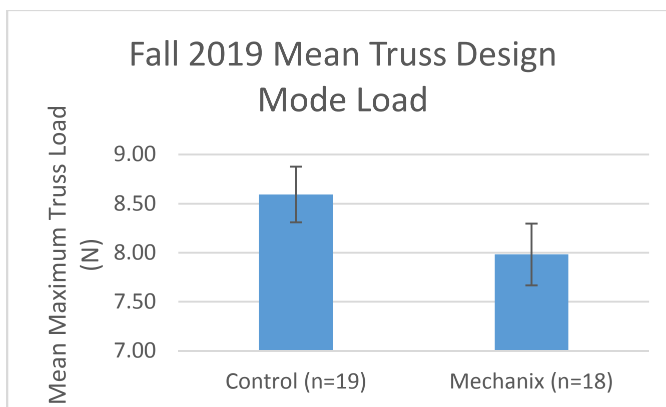

Figure 6: Fall 2019 Mean Truss Design Mode Load. The mean load applied to the student submitted truss, broken down by the experimental and control groups, highlighting additional detail.

Due to the number of students that easily came across the trivial solution, the creative design problem statement was modified to increase the difficulty, with the minimum force required being increased to 13 N. This made it so that students could not use the previous semesters trivial solution and was done with the intention of creating a more challenging environment for the students to attempt the problems. Additionally, the Mechanix application was also coded to recorded the total number of submissions a student made, this being everything from little alterations to their load answer, to drawing completely new truss systems. However, during the Spring 2020 semester the study ran into a complication, as the Covid-19 pandemic caused the education system to send students home and shift to online instruction as this assignment was due. This resulted in a decline in participation for this assignment, even after the deadline was extended by one week due to Covid-19.

During the weeks following the initial COVID-19 outbreak, students were directed to continue working on the assignment. Members of the control group were told to scan their work and submit the scans online for grading, while the experimental group was told to continue their work on Mechanix. Out of the 52 students enrolled in the class, only a total 29 total submissions were made, with 12 of those submitted assignments being from students in the control group (possible total n=27) and the other 17 coming from the experimental group (possible total n=25). Upon completing a chi-square test, it is apparent that this difference is mostly due to random chance, as the calculated P-value (0.0874) shows that the resulting difference in participation rate can be attributed to random differences in motivation.

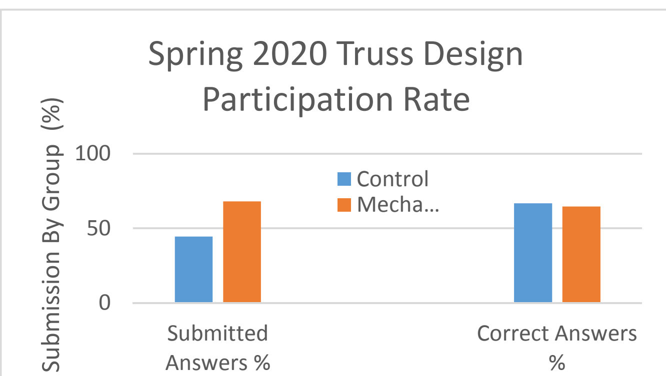

Figure 7: Spring 2020 Truss Design Mode Participation Rate. A graphical summary of the participation rate of students that completed the Truss Design mode homework. Due to the Covid-19 Pandemic, class participation on this assignment dropped significantly, resulting in more submissions using Mechanix than through the paper method. "Correct" is taken as getting a total grade greater than 90 on the assignment.

During the spring 2020 semester, despite directing students in the control group to submit all work, the total number of submitted answers did not change. Most students assigned to the control condition submitted their final answer only. Meanwhile, Mechanix automatically recorded all attempts at the problem, and once certain outlier had been removed, the mean total number of submissions on Mechanix was 10.35 submissions per user, while the control group provided an mean of 1.66 submissions per user, as seen in Figure 8. Specific students were considered outliers if they submitted an excessive amount of answers when compared to other students. This difference could be due to Mechanix recording all submissions from a user not just new iterations, but it does show that students that used the software had more engagement with the problem set on average than students that completed the problems by hand. Whatever the reason, a two-tail t-test assuming unequal variance was run on the submissions, with the results showing that there is a statistical difference between the submissions of the control and experimental groups.

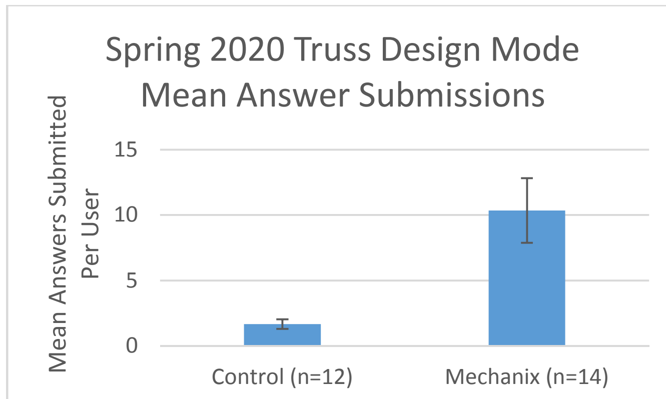

Figure 8: Spring 2020 Truss Design Mode Answer Submissions Per User. A comparison of the total number of the mean total answers submitted per user on the Truss Design mode, showing that the Experimental group submitted a statistically different number of submissions than the Control group

Because of the increased challenge in the problem statement for the Spring 2020 semester, the grades did not heavily skew towards almost all students getting a good grade on the assignment. Between the control and experimental groups, the mean grade dropped from the Fall 2019 semester, resulting in mean grade of 79% and 88% for the control and experimental groups respectably (Figures 9-10). Upon doing a two-tail t-test (t-value=-3.47703, dof=14 P=0.0037) assuming unequal variance, there is no statistically significant difference between the two groups, however the results do show a trend toward the experimental group performing better than the control group , which is consistent with past results of other homework assignments given using Mechanix [13-17].

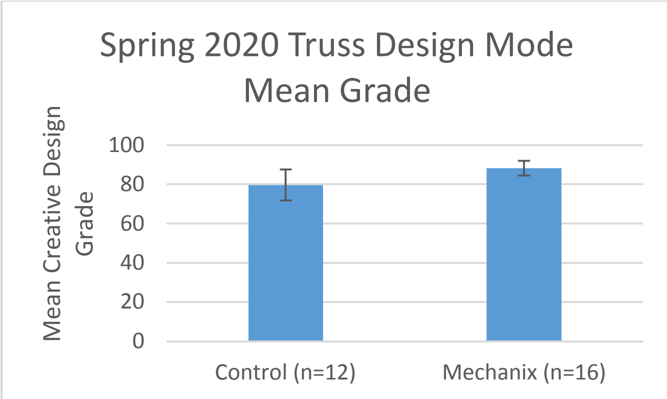

Figure 9: Spring 2020 Truss Design Mode Mean Grade. A comparison between the mean grade of the experimental and control groups on the Truss Design Mode, showing a positive trend towards the experimental group outperforming on the assignment.

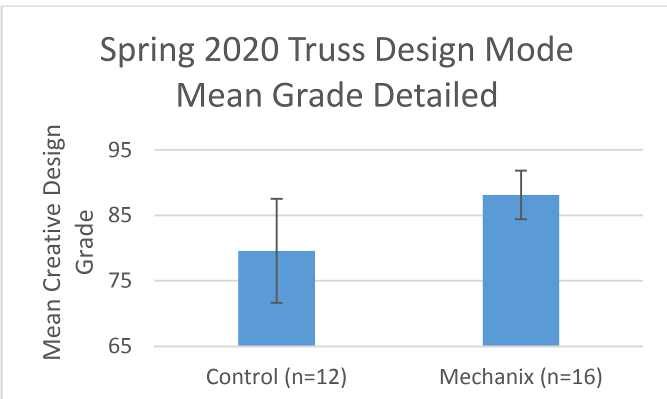

Figure 10: Spring 2020 Design Mean Grade. A comparison between the mean grade of the experimental and control groups on the Truss Design Mode, showing a positive trend towards the experimental group outperforming on the assignment. Zoomed in to emphasize the details.

One final aspect of the data that can be reported is the distinction between the mean loads that the trusses of the experimental and control groups can take. Again, the trend of the experimental group outperforming the control group continues, due to the reduction in sample due to the lack of correct answers. On the difficult problem statement, both groups on average created solutions to the design problem which exceeded the minimum truss requirements, with Figure 11 showing the small distinction between these two groups. The Control group mean a permissible load of 14.41 N, while the experimental group mean a permissible load of 14.81 N. It should be noted that one of the members of the experimental group created a truss which almost doubled the required load, holding 25.37 N. A two-sample t-test assuming unequal variance was conducted, with no statistical difference existing in the data (t-value=0.316, dof=13 P=0.7567), despite a trend towards the experimental group performing better than the control group.

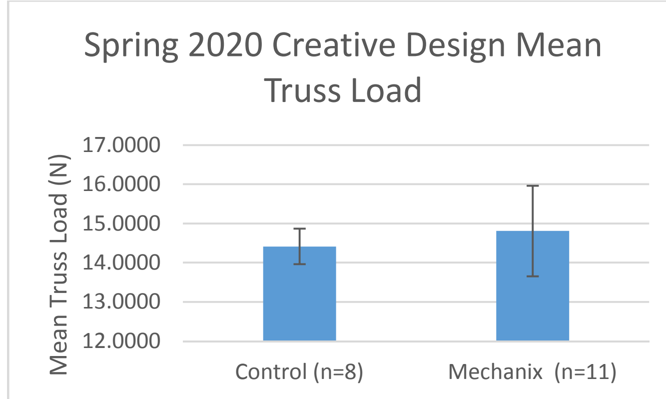

Figure 11: Spring 2020 Tech Truss Design Mode Mean Truss Load. A comparison between the mean calculated load the submitted trusses could handle, showing that the spring 2020 Variation of the problem was solvable and that the experimental group trended towards having better results.

VI. Conclusion

To better develop creative, innovative engineers, students need more opportunities to practice these skills throughout the undergraduate curriculum. Open-ended design problems are especially challenging in large, foundational engineering courses like Statics and Dynamics. Mechanix has been developed to allow for open-ended truss design problems to be naturally hand sketch by the students and then automatically graded. This study aims to evaluate Mechanix's ability to provide real-time feedback on the truss design problems and its influence on student behavior. The application is intended to provide additional details on the accuracy of a drawn FBD and providing feedback on the correctness of the user inputted solution, over traditional homework methods. This application has been used to create a powerful interactive body for the creation and distribution of challenging problems, which would normally be difficult to grade for a large class. As stated before, students that use the software to complete difficult design-based problems are seemingly more involved than their counterparts who complete these problems using a traditional, paper-based, homework method. This can be seen through the amount of submissions that the students made, as well as the tendency for the participants who completed the assignment using the Mechanix software to provide solutions that can take larger loads. It is believed that the ability to quickly input solutions, and iterate through these solutions, while getting real-time feedback allows students to explore more of the solution space than those that would complete the assignment on paper.

Through increasing the challenge of the problem, it proves that the simpler problem did not encourage the students to explore more of the design space, which can be seen by the increase in total mean calculated load in both the experimental and control groups. The data that supports this is only showing a trend however, as the Spring 2020 data's depth was damaged by the lack of participation that occurred during the global health crisis.

Going forward, future uses of the Truss Design Mode will continue to use the more difficult question variation, as to increase the sample size, as well as validate the importance of the ability of the application enabling students to explore a more difficult design space. Additionally, the Truss Design Mode is planned to be implemented at other universities, enabling a larger demographic of students to be.

Based on the results of the Truss Design Mode in both the simplified and more difficult form, a trend is observed that Mechanix allows a student to explore more of the design space, both based on the difference in grades, as well as the significant difference in submitted responses to the problem. Students using Mechanix discover more solutions to the problem, enabling more optimization of the solutions. This is reflected in both the mean total number of submissions per user, as well as the maximum calculated permissible load favoring the experimental group. More research is needed to understand the impact of Mechanix and of integrating open-ended design problems in the basic engineering classes.

VII. Acknowledgements

Partial support for this work was provided by the National Science Foundation's Improving Undergraduate STEM Education (IUSE) program under Award No. 1725423 and 1726306. Any opinions, findings, and conclusions or recommendations expressed in this material of those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] B. S. Bloom and D. R. Krathwohl, "Taxonomy of educational objectives. Handbook I: Cognitive domain," 1966.

[2] "Educating the engineer of 2020: Adapting engineering education to the new century," National Academy of Engineering, 2005.

[3] N. A. o. Engineering, *The Engineer of 2020: Visions of Engineering in the New Century*. Washington, D. C.: The National Academies Press, 2002.

[4] M. G. Sutton and C. Jong, "A truss analyzer for enriching the learning experience of students," *age,* vol. 5, p. 1, 2000.

[5] A. Giovanni, "Bridge Architect," 1/26/2014. [Online]. Available: https://www.microsoft.com/en-us/p/bridge-architect/9wzdncrfhxqh?activetab=pivot:overviewtab .

[6] R. J. Roselli *et al.*, "Integration of an interactive free body diagram assistant with a courseware authoring package and an experimental learning management system," in *proceedings of the 2003 American Society for engineering education annual conference & exposition, Nashville, USA*, 2003, pp. 1-10.

[7] J. J. Rimoli. "Truss Me!" http://www.scientificmonkey.com/software.html (accessed April 2020.

[8] W. Lee, R. de Silva, E. J. Peterson, R. C. Calfee, and T. F. Stahovich, "Newton's Pen: A pen-based tutoring system for statics," *Computers & Graphics,* vol. 32, no. 5, pp. 511-524, 2008.

[9] R. L. Bangert-Drowns, C.-L. C. Kulik, J. A. Kulik, and M. Morgan, "The instructional effect of feedback in test-like events," *Review of educational research,* vol. 61, no. 2, pp. 213-238, 1991.

[10] P. Black and D. Wiliam, "Assessment and classroom learning," *Assessment in Education: principles, policy & practice,* vol. 5, no. 1, pp. 7-74, 1998.

[11] J. A. Kulik and C.-L. C. Kulik, "Timing of feedback and verbal learning," *Review of educational research,* vol. 58, no. 1, pp. 79-97, 1988.

[12] G. Gibbs and C. Simpson, "Conditions under which assessment supports students' learning," *Learning and teaching in higher education,* no. 1, pp. 3-31, 2005.

[13] O. Atilola, M. Field, E. McTigue, T. Hammond, and J. Linsey, "Evaluation of a natural sketch interface for truss FBDs and analysis," in *2011 Frontiers in Education Conference (FIE)*, 2011: IEEE, pp. S2E-1-S2E-6.

[14] O. Atilola, M. Field, E. McTigue, T. Hammond, and J. Linsey, "Mechanix: a sketch recognition truss tutoring system," in *ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 2011: American Society of Mechanical Engineers Digital Collection, pp. 645-654.

[15] O. Atilola, E. M. McTigue, T. Hammond, and J. Linsey, "Mechanix: evaluating the effectiveness of a sketch recognition truss tutoring program against other truss programs," in *120th American Society for Engineering Education Annual Conference & Exposition (ASEE), ASEE, Atlanta*, 2013, pp. 23-26.

[16] O. Atilola, C. Osterman, T. Hammond, and J. Linsey, "Mechanix: the development of a sketch recognition truss tutoring system," *Proceedings of the American,* 2012.

[17] O. Atilola *et al.*, "Mechanix: A natural sketch interface tool for teaching truss analysis and free-body diagrams," *AI EDAM,* vol. 28, no. 2, pp. 169-192, 2014.