12-2023

Student Reliance on Simulations: The Extent That Engineering Students Rely on the Outcomes of Their Simulations

Jonathan Anderson
Utah State University, jonathan.d.anderson@usu.edu

STUDENT RELIANCE ON SIMULATIONS: THE EXTENT THAT ENGINEERING STUDENTS RELY ON THE OUTCOMES OF THEIR SIMULATIONS

by

Jonathan Anderson

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Engineering Education

Approved:

______________________________ ______________________________

Wade Goodridge, Ph.D.
Major Professor

Oenardi Lawanto, Ph.D.
Committee Member

______________________________ ______________________________

Idalis Villanueva, Ph.D.
Committee Member

Ning Fang, Ph.D.
Committee Member

______________________________ ______________________________

Don Cripps, Ph.D.
Committee Member

D. Richard Cutler, Ph.D.
Vice Provost of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2023

ABSTRACT

Student Reliance on Simulations:

The Extent that Engineering Students Rely on the Outcomes of Their Simulations

by

Jonathan David Anderson, Doctor of Philosophy

Utah State University

Major Professor: Wade Goodridge, Ph.D.
Department: Engineering Education

The purpose of this research was to investigate the factors that contributed to engineering education students' reliance on technology while learning new concepts. The researcher hypothesized that students would give reliance to their technology, even in the face of evidence that the technology was not working as intended. This research used a mixed-methods approach to answer the research questions. Three questions guided the research: (1) How are the participant's level of automation complacency and the correctness of the simulation that participant is using related?; (2) How is automation bias related to a participant's ability to recognize errors in a simulation?; and (3) What factors explain the automation bias and automation complacency that the participants are experiencing? The third research question had two subquestions: (a) What factors explain the correlation between a participant's level of automation complacency and the correctness of the simulation that participant is using?; and (b) What factors explain the impact that automation bias has on a participant's ability to recognize errors in that simulation?

This study was based on the Theory of Technology Dominance, which states that people are more likely to rely on their technology the less experience they have with the task, the higher the complexity of the task, the lower the familiarity with the technology, and the further the technology is from the skillsets needed to solve the problem. This framework is built on the automation bias and automation complacency given by an individual towards technology. Automation bias is an overreliance on automation results despite contradictory information being produced by humans, while automation complacency is the acceptance of results from automation because of an unjustifiable assumption that the automation is working satisfactorily.

To ensure that the study could gather the information necessary, the mixed-study utilized deception techniques to divide participants into separate groupings. Four groupings were created, with some participants being given a properly functioning simulation with others being given a faulty simulation. Half of each grouping were informed that the simulation may have errors, while others were not. All participants who completed the study were debriefed about the real purpose of the study, but only after information had been gathered for analysis. The simulation given to all participants was designed to help students learn and practice the Method of Joints.

Students participating in the statics courses taught in the College of Engineering courses at Utah State University were invited to participate in the program over Spring and Fall semesters of 2022. Sixty-nine participants began the study, but only thirty-four remained in the study through to completion. Each participant took a pre-questionnaire, worked with a provided simulation that was either correct or incorrect, were possibly informed of potential errors in the simulation, and took a post-questionnaire. A few participants were invited to participate in an interview.

The findings of this study revealed that students often have high levels of automation bias and automation complacency. Participants changed their answers from wrong answers to right answers more often when using correct simulations and from right answer to wrong answers more often when using faulty simulations. The accuracy of each participant's responses was also higher for those with correct simulations than faulty simulations. And most participants expressed that they checked their work and changed their answers when the simulation asked them to. These findings were confirmed through the use of the post-questionnaire results and in interview analysis between the groups.

(292 pages)

PUBLIC ABSTRACT

Student Reliance on Simulations:

The Extent that Engineering Students Rely on the Outcomes of Their Simulations

by

Jonathan David Anderson

The purpose of this research was to investigate the factors that contributed to engineering education students' reliance on technology while learning new concepts. The researcher hypothesized that students would give reliance to their technology, even in the face of evidence that the technology was not working as intended. This research used a mixed-methods approach to answer the research questions. Three questions guided the research: (1) How are the participant's level of automation complacency and the correctness of the simulation that participant is using related?; (2) How is automation bias related to a participant's ability to recognize errors in a simulation?; and (3) What factors explain the automation bias and automation complacency that the participants are experiencing? The third research question had two subquestions: (a) What factors explain the correlation between a participant's level of automation complacency and the correctness of the simulation that participant is using?; and (b) What factors explain the impact that automation bias has on a participant's ability to recognize errors in that simulation?

This study was based on the Theory of Technology Dominance, which states that people are more likely to rely on their technology the less experience they have with the task, the higher the complexity of the task, the lower the familiarity with the technology, and the further the technology is from the skillsets needed to solve the problem. This framework is built on the automation bias and automation complacency given by an individual towards technology.

Automation bias is an overreliance on automation results despite contradictory information being produced by humans, while automation complacency is the acceptance of results from automation because of an unjustifiable assumption that the automation is working satisfactorily.

To ensure that the study could gather the information necessary, the mixed-study utilized deception techniques to divide participants into separate groupings. Four groupings were created, with some participants being given a properly functioning simulation with others being given a faulty simulation. Half of each grouping were informed that the simulation may have errors, while others were not. All participants who completed the study were debriefed about the real purpose of the study, but only after information had been gathered for analysis. The simulation given to all participants was designed to help students learn and practice the Method of Joints.

Students participating in the statics courses taught in the College of Engineering courses at Utah State University were invited to participate in the program over Spring and Fall semesters of 2022. Sixty-nine participants began the study, but only thirty-four remained in the study through to completion. Each participant took a pre-questionnaire, worked with a provided simulation that was either correct or incorrect, were possibly informed of potential errors in the simulation, and took a post-questionnaire. A few participants were invited to participate in an interview.

The findings of this study revealed that students often have high levels of automation bias and automation complacency. Participants changed their answers from wrong answers to right answers more often when using correct simulations and from right answer to wrong answers more often when using faulty simulations. The accuracy of each participant’s responses was also higher for those with correct simulations than faulty simulations. And most participants expressed that they checked their work and changed their answers when the simulation asked them to. These findings were confirmed through the use of the post-questionnaire results and in interview analysis between the groups.

DEDICATION

This work is dedicated to my beloved wife, Trenna, for her support of my studies and work during times of doing my dissertation and to my beautiful children, DJ and Hailey.

This work is also dedicated to my mother, Lujean, my father, John, and my step-mother, Alison, for their prayers, support, and encouragement throughout all of my studies.

I dedicate this work to my mother-in-law, Debbie, and father-in-law, Gary, for the support and encouragement they have provided, as well.

Finally, I dedicate this work to Dr. DJ Lee, my first graduate advisor, who has been a faithful and supportive friend through all of my education.

ACKNOWLEDGMENTS

I would first like to acknowledge my Heavenly Father and the hand He has played in my life, by giving me the family and experiences of my youth and the abilities and knowledge that have been instrumental in my pursuit of an education and a career in higher education. Although there have been many struggles during this journey, I know that He has been supportive of me and has blessed me greatly.

I would also like to acknowledge those who have been so supportive of me and have helped me complete this work. I specifically need to thank Wade Goodridge, my major advisor, for his support and encouragement during my doctoral study. I cannot thank him enough for both giving me the chance to pursue this degree or the support he personally provided during it. His insights have been helpful in many aspects of my life.

I am also thankful to my committee members, Dr. Oenardi Lawanto, Dr. Idalis Villanueva, Dr. Ning Fang, and Dr. Don Cripps. Each of them contributed directly to my research, providing guidance and helping me to push the boundaries of my study. All of them provided vital and necessary suggestions that made this study better than it would have been without them.

Additional thanks must be given to my Department Chairs in the Department of Mathematical and Quantitative Reasoning at Utah Valley University, Keith White and Evelyn Porter. I could not have pursued this degree without the support they provided to me in selecting a compatible schedule to continue to teach full-time while pursuing the doctoral degree full-time, as well.

I also would not have continued to pursue this degree without the support of Dr. DJ Lee of Brigham Young University, who was the first graduate advisor I worked with in my academic career. His encouragement to not give up on a doctoral degree and his support when I applied at Utah State University cannot be understated. I am forever grateful for the support and guidance he has continued to provide to me.

Finally, I would like to thank my family. I could not have done this without each of you who have allowed me to take the time to pursue this degree. Thank you for your sacrifice in letting your husband and father take the time to accomplish this degree. I love you all!

Jonathan David Anderson

CONTENT

LIST OF TABLES

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

Background of Study

We live in a world where technology is advancing rapidly (Bryant, 2011). Much of this technology is aimed at making our lives easier. Engineering students are often provided with tools that both aid in understanding and solving their problems, ranging from calculators to sophisticated simulation software. But students may be relying too much on the results of these tools, using them as a crutch to learning the fundamental principles found in the problems that the tool is helping the student to solve. Do simulations impede or accentuate student's conceptual/analytical knowledge acquisition?

Given the recent challenges in the world, such as Covid-19, the need for appropriate technologies has become even more apparent (National Science Foundation, 2021). While people across the world were quarantined, many relied on technology to continue to work or to engage in learning activities. The National Science Foundation (NSF) (2021), calls for researchers to "transform learning through innovative teaching, educating, and mentoring practices in a variety of settings, to include formal settings such as physical and virtual classrooms, as well as informal settings (e.g., museums, nature centers, libraries, citizen-science activities, and other on-line experiences)" (p. 4). Simulations used to help students learn appropriate concepts fall within the scope of this transformation, but only so far as they are built with appropriate "expertise across disciplines, including learning sciences, discipline-based education research, computer science, engineering, human-computer interaction, design, social and behavioral sciences together with ethics, policy, and privacy" (NSF, 2021, p. 4).

Simulations fall into three broad categories (Magana et al., 2012):

1. Automated Simulations: Students watch the simulation with little to no interaction. Often used as an alternative to lectures.
2. Computer Simulations: Students use an interface to interact with the simulation. Computer simulations are often built for specific learning objectives.
3. Computational Simulations: Students use a broad computational engine, typically one used by practicing engineers.

While Computational Simulations often mirror the real environment that engineers may find themselves in, Computer Simulations see far more use in the educational world. This is partly because they are easier to build, as they focus only on a single learning objective, but they are also typically easier for students to engage with (Magana et al., 2012). Computational Simulations require a deeper understanding of the environment of the simulation, which adds to the complexity of their use, including the potential for a prelab lecture to help students be successful in using the computational engine (Brophy et al., 2013). Automated Simulations are usually not studied, as they lack the capability of user interaction, limiting their use to being a supplement to or replacement of traditional lectures (Magana et al., 2012).

Early research in simulations involved using simulation games to explore concepts in the social sciences (Dukes & Waller, 1976). Early simulations did not use computer technology. Instead, physical models were simulated, sometimes with cardboard boxes and strings (Agnew & Shinn, 1990). Computer simulations became more widespread in the mid-1990s (Nahvi, 1996). Research in this area began to focus on the different ways to present information, and user interfaces were often a large weakness of simulations from this time period. As computers became more powerful, several companies saw opportunities to develop specialized computer simulations. Wankat (2002) tested several commercial simulations in a traditional lecture-oriented class and found that students performed better on unit exams when they had access to simulations.

Simulations began to spread and became more prevalent in education. Some of these simulations were directly integrated with learning material for classes, but most remained separated from the learning material (van der Meij & de Jong, 2006). Simulations became more complicated. Some simulations focused on emulating specific high-end equipment, allowing students a safe place to use equipment without the potential of damaging said equipment (Koh et al., 2010). The quality of the simulation also impacts the student interaction, which can affect the outcomes – good design is crucial in having a functional simulation (Nahvi, 1996).

Computational simulations have started to see more use starting in the early 2010s (Magana et al., 2012; Uribe et al., 2016). These simulations allow more work to be done, as the simulation can be directly interacted with by the students. These students also gain real-world experience using real tools from industry, such as the nanotechnology simulator found at nanoHUB.org (Brophy et al., 2013; Magana et al., 2012; Uribe et al., 2016).

One of the inherent risks in technology is that people can become overly reliant on that technology (Grissinger, 2019; Lou & Sun, 2021). Two different forms of overreliance on technology have been identified by Billings et al. (1976): automation bias and automation complacency. Automation bias is the overreliance on automation in accomplishing a task while ignoring contradictory information produced by humans. Automation complacency is a sense of self-satisfaction that may result in a lack of vigilance due to the unjustified assumption that the automation is in a satisfactory system state. That is, automation complacency is the trust that the automation is performing as it is supposed to without any verification that it is doing so.

Both can play a role in the way that students interact with simulations. A student who has automation bias will accept the results of the simulation since the simulation is built to handle the situation, ignoring their own insights and thoughts as the computer "cannot be wrong." Similarly, a student who has automation complacency assumes that the simulation is in a correct state to work, thus furthering their belief that the simulation cannot be wrong. Because of the prevalence of these two measures of overreliance on technology and automation, students often immediately trust the results of their simulation. It is only after the homework, quiz, or exam score is returned that students find that their faith in the technology may not have been well founded, but, potentially in line with the automation bias that they hold, the student may still blame either themselves or their grader rather than analyzing the use of the simulation.

The Theory of Technology Dominance helps to explain these discrepancies (Arnold & Sutton, 1998). In this theory, there are four key factors: (a) Task experience, which is the level of experience accumulated by a user to complete a certain task; (b) Task complexity, which is the degree of cognitive abilities needed to complete the task; (c) Decision aid familiarity, which is the user's familiarity with the automated tool in question; and (d) Cognitive fit, which is how well the automated tool matches the skillset to solve the problem without it. The Theory of Technology Dominance suggests that a user is more likely to over-rely on a technology if they fall into one of the following categories: (a) a user who has a low level of task experience, and (b), a user who has high levels in all four factors. While students often fall immediately into the first category, the second category is interesting. This category includes faculty who rely on aids, such as the Web of Science, to help them in their literature reviews (Lou & Sun, 2021). The very fact that they often use the tool to help them may make experienced users less likely to notice errors in the tool when it excludes papers that would otherwise be included in a search of the database. For purposes of this study, the researcher is more interested in the first category, as students are much more likely to have low task experience and potentially low decision aid familiarity.

Purpose and Objectives

The purpose of this research is to investigate the extent to which students show automation bias and automation complacency when working with a simulation. The researcher hypothesized that students had high levels of automation bias and automation complacency since they will have low levels of task experience and decision aid familiarity, as they are working through new material and using a new system in the simulation. The simulation utilized directions and calculations to help students learn a specific process. The proposed work follows an explanatory multi-method model.

The following objectives framed and guided the research:

1. To investigate the level of automation complacency shown by students as a measure of whether students verified the results a truss simulation led them to.
2. To investigate the level of automation bias shown by students as a measure of their task experience in solving truss problems with the method of joints.
3. To develop an understanding of why students may exhibit different levels of task experience, automation bias, and automation complacency.

Research Questions

Previous research in simulations have tried to show that simulations improve student learning (Bing et al., 2014; Brophy et al., 2013; Chaturvedi et al., 2011; Chyung et al., 2010; Dabbagh & Beattie, 2010; Dang et al., 2017; Fedorova et al., 2016; Gero et al., 2014; Goeser et al., 2011; Khan & Singh, 2015; Koh et al., 2010; Kollöffel & de Jong, 2013; Meschke et al., 2019; Navaee & Kang, 2017; Nikolic et al., 2011; Rojko et al., 2010; Rokooei et al., 2017). Most of these studies showed that students benefit on subsequent quiz and exam scores after working in the simulator. One of the major advantages that these studies rely on is the fact that students who used simulations were just as effective in working with the real thing as those who learned from the real thing (Agnew & Shinn, 1990). This finding was vital, as it helped encourage the use of simulations in engineering education.

Despite this research into the ability to learn from simulations, very few studies have been done on a student's interactions with the simulator. Much of the research in simulations was novel at the time, and this novelty needed to be investigated to show that students could, in fact, learn from simulations. However, this focus on the novel impact of technology on education may have potential detrimental impacts on learning if the interaction of the student and the simulation is not measured (Brown, 2009; Dalcher, 2007; Sorensen & Snider, 2001). With this target identified the following research questions are intended to be investigated.

1. How are the participant's level of automation complacency and the correctness of the simulation that participant is using related?
2. How is automation bias related to a participant's ability to recognize errors in a simulation?
3. What factors explain the automation bias and automation complacency that the participants are experiencing?
    1. Subquestion: What factors explain the correlation between a participant's level of automation complacency and the correctness of the simulation that participant is using?

2. Subquestion: What factors explain the impact that automation bias has on a participant's ability to recognize errors in that simulation?

Research Design

The research design of this study used a multi-method sequential explanatory approach to describe students' over-reliance on technology as they work through a simulation to help them better understand the conceptual knowledge. A convenience sample will be used in this study to help limit the possible confounding variables that might otherwise be present. One confounding variable was the selection of a single learning objective – the ability to use the Method of Joints to analyze a truss – which is taking place in a particular class. The research is limited in scope by this work operating under a dissertation's timeframe. For this reason, the Statics class taught at Utah State University has been selected. An advantage in using this class is that it generally has around two hundred students from many different engineering majors (civil, mechanical, etc.), which will help to diversify the population in the study. A convenience sample can be appropriate in situation such as this, as there are naturally volunteers to participate in the study (Creswell & Poth, 2016).

This study was designed as a multi-method sequential explanatory design, where quantitative data is collected first, and qualitative data is gathered after to help illuminate, corroborate, or explain the findings from the quantitative study (Ivankova et al., 2006). One of the strengths of this type of study is that it is straightforward and provides opportunities to explore the quantitative results in more detail. An advantage is the deeper investigation of typical quantitative findings using qualitative methods. According to the Johns Hopkins Bloomberg School of Public Health, "mixed methods help understand, not just whether an intervention works, but how, why, and for whom" (Why Mixed Methods, n.d., pra. 5). Combining quantitative analysis with qualitative analysis can add the how, why, and whom to the results, making them richer. In addition, qualitative research can shed insight even when randomized controlled trials suggest that a particular intervention has failed (Trevas & Nimkoff, 2015). This insight may allow future research to continue working on the intervention, allowing for better instruction and interaction with participants and providing a better foundation for future studies. Similarly, using a single method can often lead to a limited interpretation of data (Trevas & Nimkoff, 2015). However, it can often take a long time to gather all the results, as it can be difficult to collect and analyze both types of data. The timeline for this research is shown in Figure 1. The timeline shows the gathering of Quantitative data prior to the gathering of Qualitative data, showing the supporting work that the Qualitative data would provide to the Quantitative data found.

Figure 1

Visual model for mixed methods design procedures

Activity	QUAN Data Collection →	QUAN & qual Data Collection →	QUAN & qual Data Analysis →	qual Data Collection →	qual Data Analysis
Procedure	Demographic Questionnaire (QUAN)	Guided Simulation (QUAN) Reliance on Technology Questionnaire (QUAN & qual)	Data screening Frequencies RStudio quantitative software Thematic analysis Interpreting results	Recorded interview	Thematic analysis Interpreting results
Product	Numeric Data Assignment to Groups	CSV file with simulation results Numeric data Written responses	Descriptive statistics Frequencies and rates Paramateric and nonparametric statistics (*t*-tests, ANOVAs) Survey member checking Issues Interrater reliability Interpreted results Quantized questionnaire results	Text data (interview transcripts)	Issues Interrater reliability Interpreted results Quantized interview results
Timeline	Second month of semester	Second month of semester	After information gathered	End of third month of semester	After interviews gathered

Note. Timeline is relative to each semester, Spring 2022 and Fall 2022, separately.

Note. Capital letters in Activity indicate primary work while lower case letters represent secondary work

Due to the nature of the research questions being asked, participants were not fully informed of the full scope of the study at the beginning. That is, participants participated in a deception study (Boynton et al., 2015; Kemmelmeier et al., 2003; Sommers & Miller, 2013; Uz & Kemmelmeier, 2017). Deception studies are not seen a lot in engineering educational research and more detail about them is warranted here in the narrative. Kemmelmeier et al. (2003) outlines seven potential areas that deceptive studies have to address: Moral Turpitude, the Necessity of the Deception, the Validity of the Deception, the Violation of Participant Rights, the Potential Harm of Deception, the Potential Harm to the Discipline, and the Potential Harm to Society. The researcher advocates that deception is necessary in this case, because when a participant knows a simulation could contain errors the results of the study may be unfavorably swayed in work targeting research questions as this does. Part of what is being investigated here is the level of automation complacency and bias that the participants exhibit, and if they were informed ahead of time that the simulation may contain errors, they may have exhibited lower levels of automation bias or automation complacency than they otherwise would have. This addresses the Necessity and the Validity of the Deception. The other categories will be explored in more detail in Chapter III.

Participants were divided into four groups along two axes. One axis divided groups into those who know there are potential errors in the simulation and those who do not. The second axis divided groups into those who experienced a simulation with errors and those who experienced the simulation with no errors. This last grouping covered participants receiving correct instructions in their guided simulation and the others who will not have these correct instructions. This provided insight into how students handled the simulation, which shed light on the amount of automation bias and complacency that may arise. These groups are shown in Table

1. Students were divided evenly between these four groups as they completed the pre-questionnaire, starting with the first of each semester who completed to remove any potential bias in treatment group placement by the researcher. That is, the first student to complete the prequestionnaire was added to the control group; the second student was added to the deceived group; the third student was added to the informed group; the fourth student was added to the misled group; the fifth student was added to the control group; and the pattern repeats from there. This setup was done independently for each semester, so the first student from each semester was in the control group.

For this study, quantitative data was gathered with an initial questionnaire including demographic data followed by results from using the simulation. These quantitative measures will inform us as to the results of the students interacting with the simulation while also not greatly impacting their grades. Participants who are in the groups containing potential errors were given the opportunity to use the simulation without the errors and will receive specific attention to help them overcome the potential harm of the simulation.

Table 1

Treatment Groups

		Simulation Errors	
		No Errors	Errors
User Knows of Potential Errors	Not Informed	Student not informed of errors, and no errors present (control) [9]	Student not informed of errors, errors present (deceived) [9]
	Informed	Student informed that errors may happen, but no errors present (informed) [8]	Student informed that errors may happen, and errors are present (misled) [8]

After the quantitative data is gathered, participants took a post questionnaire, and participants who indicated they were interested were selected for a follow-up interview. The post questionnaire started with quantitative questions, included the debriefing of the true purpose of the study, and gathered qualitative data related to how students feel about the category that they were part of. The interviews allowed the researcher to delve farther into the experiences of members from each group. The qualitative questions from the post questionnaire and the interviews were quantized to allow for analysis.

The quantitative analysis applicable to this study included descriptive, parametric, and nonparametric statistics. The qualitative data will be gathered from questionnaires, which include both Likert-scale and open-ended questions, and interview transcripts to find support for the quantitative results. Quantitative methods are most often interested in measures of central tendency, variability, associations, and significant statistics (Boone & Boone, 2012). In particular, Likert data that is only measured by which is bigger (ordinal data) is different than data which has known differences between groupings (interval data). Ordinal data uses medians, frequencies, Kendall's tau B, and chi-square tests to explore the statistics, while nominal data uses means, standard deviations, Pearson's r, ANOVAs, and t-tests to explore the statistics (Boone & Boone, 2012). These differences and which tests were used is based on the different questionnaires and will be outlined in Chapter III. Interview transcripts were be collected, coded, analyzed, and compared to the quantitative data to look for patterns that can help explain the findings. This data will be combined, looking for connections between the Likert-scale questions, the simulation use results, and the coding of individual questionnaire and interview results.

Significance of the Study

Because student reliance on the results of simulation technology has not been fully explored and the use of simulations in a Statics class is still emergent, this study seeks to develop knowledge on the impact that the Theory of Technological Dominance has on student use of simulations so that those faculty who use them to teach can either create better simulations or can offset the dangers through other means. It also demonstrates a successful deception study in engineering education.

Assumptions

Assumptions of the study are listed below:

1. *Technology Continues to Expand*: The use of technology in education is going to continue to develop and expand, changing the landscape for our new students. This includes the growing use of simulations of all types as teaching/learning aids in the classroom.
2. *Reactivity*: The researcher assumes that participants are being truthful in their responses to the questionnaires and simulation. He also assumes that the responses reflect the actual thinking that is happening at the time the participants filled out the questionnaire. He assumes that no outside influence is being exerted on the students at these times. Questionnaires are written in such a way as to not have undue pressure on the students.

Limitations

Limitations help define the scope of the research. The limitations of the study are identified in the following areas:

1. *Convenience Sample*: Due to tying into a learning objective, this study takes place within a class. The particular class, Statics, included a large number of students, which had the potential to make it easier to generalize the data and provide enough data to talk about the demographic impacts of the study, but it is still primarily a sophomore level class taken by engineering students from many disciplines, limiting the generalization. These students will have less experience with their education and with the topic being studies. In addition, all participants in the study came from Utah State University, further limiting the generalization.
2. *Quasi-Experimental*: Because the study took place in a single class, there is not a full control group to compare the results with. Some of the metrics used had controls built in, but the only students the researcher can compare to without the use of simulation results are students from previous semesters. This means that any comparison made is anecdotal at best, and may lead to generalizations which should not be fully relied upon.
3. *Seasonal Effects*: The study will took place with students in the Spring and Fall semesters, so there is minimal possibility of seasonal effects that affect the study. With future work this variable may be addressed but limitations imposed with graduate student deadlines preclude this at this point.
4. *Sample Size*: Although each semester had roughly two hundred students, only a handful participated in the study. Further, many opted out of the study. This limits the results from being fully generalizable, but the information learned is transferrable. With more samples, more institutions, and different participants, this work could become more generalizable.

5. *Self-reported Data*: The questionnaire used to gather data from the students will be self-reported.

6. *Longitudinal Effects/Timing*: This study is intended to be done with participants over a single semester with two semesters total included in the research, facilitating the graduation of a graduate student, and cannot, in this instance, be extended beyond that.

7. *Risk and Mitigation Plan*: Due to the deceptive nature of this study and the nature of explanatory mixed-methods research, it is important for the researcher to identify potential risks in the research and plan on mitigating these. The risks and mitigation of the deception is discussed in Chapter 3 under the Deception Studies topic. Issues related to potential risks in coding and in instrumentation is also discussed in Chapter 3 under Data Analysis, while risks related to handling the privacy and data of the participants is discussed in Chapter 3 under Data Handling. Specifically, harm to those students who are actively mislead in improper use of the Method of Joints will be informed of such and given the opportunity to learn using a correctly working version of the simulation; interrater agreement was measured by having additional support in coding; a factor analysis will take place on the questionnaire to ensure it is working as the original did; and participant data will be anonymized in storage, separating the identity of the student from their data. Over half of the participants which began the study opted out either by not completing the post questionnaire or by opting out after being debriefed – there is no indication to what level any of these participants took advantage of the debriefing.

8. *Researcher Biases*: One potential bias that might be present is that the graduate researcher is very supportive of using technology in education, which could have potentially influenced the interpretation of results. Recognition and tempering of this

potential bias was monitored by the course instructor but questions are asked in the study to additionally address any bias here. Specifically, the demographic questionnaire asks questions about prior experience using learning programs, while the Reliance on Technology questionnaire has questions about the participant's beliefs about using simulations to learn and provides them the ability to provide feedback specifically about what they would change with the program. They are also asked about the process after finding out that the system may have contained errors. These themes are also part of the interview protocol, with questions about using guided simulations and a chance to discuss the deception that has occurred. In these ways, it is hoped that the participants were able to explain their thoughts without the bias of the researcher interfering.

Delimitations

Like limitations, delimitations are specifically selected to define the scope of the research. The delimitations of this study are:

1. *Selection of Statics Class at Utah State University*: Students will come from many different engineering programs, helping to make the results more generalizable, but they will still be missing a few (Electrical and Computer engineering students will be completely missed, for example). This is done to make sure that the simulation specifically aligns with a learning objective. The class itself is also a larger class, which allowed for students representing a more diverse background to participate. However, by selecting a single specific school and its college, it is more likely that the group will be more homogenous than desired.
2. *Method Taught in Simulation*: The Method of Joints has been selected as the topic that the simulation is being built around. This method is chosen because students encounter

the method earlier and it is a more direct method that the simulation can easily solve and prepare for. In building the simulation, specific problems will be created by the researcher rather than being built to handle any/all joint problems. This is done partly to constrain what students will be working on and also to ensure that it is used as a learning tool and not a homework aid tool - future studies and development may be done to broaden this application in the future (both as a homework aid tool and in expanding the methods that can be used in the simulation).

3. *Introduction of Errors*: Specific errors that are commonly made by students were included in the simulation, thus emulating the types of errors that students would commonly see.
4. *Random Sampling*: The participants were placed into one of the single control group or three treatment groups. Participants were randomly assigned by being placed into groups in the order in which they completed the prequestionnaire at the start of the study for the quantitative portion of the research.
5. *Longitudinal Effects/Timing*: This study is intended to be done over a single semester for the participants and over two semesters for the researcher, facilitating the graduation of a graduate student, and cannot, in this instance, be extended beyond that.

Definition of Key Terms

The definitions of the terms or phrases below are for clarification and understanding with reference to this study.

1. *Automation Bias:* An overreliance on automation results while potentially ignoring contradictory information produced by humans (Billings et al., 1976).

2. *Automation Complacency:* The acceptance of results from an automation because the user unjustifiably assumes that the automation is working satisfactorily (Billings et al., 1976).
3. *Reliance:* A measure of how much a participant relies on the results of a simulation. Tied directly to automation bias and automation complacency, previously defined, and the Theory of Technology Dominance, by suggesting why users trust the results of automated processes (Lou & Sun, 2021).
4. *Traditional Student*: A student who starts working towards a degree at the customary age, generally right after high school (Bean & Metzner, 1985).
5. *Nontraditional Student*: A student who is "older than 24, or does not live in a campus residence, or is a part-time student, or some combination of these factors; is not greatly influenced by the social environment of the institution; and is chiefly concerned with the institution's academic offerings" (Bean & Metzner, 1985). Includes traditional students who step away from pursuing their degree for more than three years and generally have to begin again.
6. *Simulation*: The production of a computer model of something, typically used to help students learn a particular concept (Magana, 2017).
7. *Computer Simulation*: A type of simulation specifically intended to help students understand a topic. Generally includes a user interface and controls that allow the student to modify the simulation to see results. Also, not generally a tool used by those in the industry (Magana, 2017).
8. *Convenience Sample*: The sample of participants is taken from a convenient place, such as a course.

9. *Strategic Sampling:* A sampling method which is carefully planned but otherwise random (Li et al., 2013). Examples of strategic sampling include dividing a class into high and low performers to investigate research questions.

10. *Statics*: A study of engineering mechanics. Typically the first class in this area where curriculum focuses on forces acting on a rigid body at rest or moving at a constant velocity.

11. *Truss Analysis (Method of Joints)*: Analytical technique to isolate forces acting in tension or compression within a truss member using the derivation and solving of equilibrium equations. The point of focus is the joint or connection between members.

12. *Case Study*: The case study method "explores a real-life, contemporary bounded system (a case) or multiple bounded systems (cases) over time, through detailed, in-depth data collection involving multiple sources of information… and reports a case description and case themes" (Creswell & Poth, 2016).

13. *Deception Study:* A study where the researcher deliberately creates a situation in which the participant's beliefs about the situation are different from the knowledge that the researcher has of the same situation (Kemmelmeier et al., 2003).

Organization of this Dissertation

The dissertation is organized as follows. Chapter I focuses on an introduction, background, objectives, and design of the study. Chapter II is a review of the literature related to reliance, simulations, constructivism, and deception studies. Chapter III reviews the Research Methodology, including the methods for data collection, the study participants, and data analysis. Chapter IV shows the findings after evaluating the data. Chapter V discusses the findings and their impact on the research questions and indicates potential areas of further study.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

Technology continues to grow and expand, giving us ever increasing ways to engage in the environment around us (Bryant, 2011; Sorensen & Snider, 2001). While much of this technology provides society a more comfortable living, it also carries inherent dangers. Some of these dangers relate to the technology failing, while others relate to the humans who use that technology (Bryant, 2011; Clubb, 2010; Sorensen & Snider, 2001). As people start to use more and more technology, they may begin to become overly reliant on the technology, trusting that it is working as intended (Lou & Sun, 2021).

The purpose of this literature review was to synthesize current and historical literature on the topics of reliance on technology, proper use of simulations in education, the state of research on the use of simulations, and the use of deception studies in research. The objectives of this review were to:

1. Describe the current state of research on the reliance on technology.
2. Discuss the current state of research regarding the use of simulations in engineering education.
3. Discuss the current use of simulations in education, particularly in the field of engineering education.
4. Discuss the use of deception in research in education.

The use of simulations and technology in education has been discussed in many different domains of learning, including in education, medicine, engineering, transportation, and aerospace, as shown in Table 2. The following databases were searched: EBSCO, IEEE Xplore, Google Scholar, and Science Direct. The journals and conference proceedings found show that there is a growing interest in the study of simulations, such as *IEEE Education Engineering Conference, Advances in Engineering Education, ASEE Annual Conference, Educational Technology and Society, IEEE Transactions on Education,* and the *Journal of Engineering Education*. The following key words and terms were used to obtain this body of literature: reliance on technology, simulation, computer simulation, simulation-based learning, virtual labs, computational simulations, remote laboratory, and deception.

The key words formed the first level of inclusion criteria – once papers were shown to contain these words in either their title or key words, they were reviewed to see how applicable they were to the study at hand. Papers were divided according to the above objectives. For objectives 1, 3, and 4, papers were excluded if they were too narrow in focus, such as on a particular aspect of simulations rather than being about simulations in general. For objective 4, papers that were about deception itself, not its use in research, were also excluded from the work. For objective 2, papers were included if they discussed using simulations in learning and were excluded or moved to objective 3 if they did not actually measure the use of the simulation in learning. A few papers were found to relate to multiple objectives, particularly between objectives 2 and 3. A total of eighty-five (85) papers were found during this literature review, of which twenty-three (23) were excluded. There were eighteen (18) papers included about reliance on technology (objective 1), eleven (11) papers included about the use of simulations, twenty-eight (28) papers included about the use of specific simulations (also shown in Table 2), and five (5) papers found about the use of deception in research studies.

Table 2

Areas Using Simulations

Domain of Learning	Related Articles	Uses of simulations
Education	Deley and Dubois (2020) Lou and Sun (2021) Sorensen and Snider (2001)	Literature reviews, social media networking, market data
Medicine	Campbell et al. (2007) Grissinger (2019)	Healthcare data entry, patient diagnosis, prescriptions
Engineering	Agnew and Shinn (1990) Bing et al. (2014) Brophy et al. (2013) Chaturvedi et al. (2011) Chyung et al. (2010) Dabbagh and Beattie (2010) Fedorova et al. (2016) Khan and Singh (2015) Koh et al. (2010) Kollöffel and de Jong (2013) Meschke et al. (2019) Navaee and Kang (2017) Nikolic et al. (2011) Rojko et al. (2010) Rokooei et al. (2017) Shao et al. (2008) Tang (2014)	Building and material reactions to outside stimuli, materials, virtual laboratories, machining, electrical circuits, tunneling, construction, microcontrollers
Transportation	Grant et al. (2009) Kos et al. (2013) McCullough and Collins (2019)	Global positioning system, navigation, smartphone response to weather conditions
Aerospace	Bryant Jr (2011) Clubb (2010) Dalcher (2007)	Sensors, flight

Reliance on Technology

While technology is constantly developing and expanding in capability and ability, this study is most interested in technology growth in education. The main goal of technology in education is to serve the needs of the learner, whether they are in physical classes or online (Sorensen & Snider, 2001). There is an inherent danger of using technology simply because it is new rather than choosing to use it for pedagogical purposes. Technology should be selected because it meets learning goals rather than for the sake of being used (Brown, 2009). Brown (2009) specifically argues that technology should not replace the personal interactions that take place in a class. Doing so actually diminishes the value of education. As technology expands, it is possible that its users, specifically those new to the technology and what it helps solve, begin to over-rely on the technology, leading them to making bad decisions or failing to learn the appropriate concepts.

The Theory of Technology Dominance, proposed by Arnold and Sutton (1998), outlines several interrelated factors that can help explain this overreliance on technology. The first factor, Task Experience, is related to the user's knowledge of the task the technology is being used to accomplish. High levels of Task Experience show that the user is very familiar with the process and can likely perform it on their own without the technology. Low levels of Task Experience show that the user is unfamiliar with the process and is unlikely to perform the task satisfactorily without help.

The second, Task Complexity, is determined by the degree of cognitive abilities needed to complete the task. Complex tasks that require active thought will have high levels of Task Complexity, while tasks that can be accomplished quickly and with little thought will have low Task Complexity.

The third, Decision Aid Familiarity, assumes that the technology is helping the user make a decision, but it still can be applied to other tools. This concept is related to the level of familiarity that the user has with using the technology. The more experience they have, the higher this concept is rated.

The final factor, Cognitive Fit, is a measure of how well the technology matches the skillsets needed to solve the task without it. The closer the technology matches the skillsets, the higher the Cognitive Fit. This is often one of the more difficult aspects of the theory to measure but can often be done by an expert in the task analyzing the technology in use.

One of the fascinating outcomes of the theory is that those who are most likely to overly rely on technology are those who either have low Task Experience or those who have high levels in all four factors. The first makes a lot of sense, as the researcher would expect those unfamiliar with a task to have low confidence in their abilities, forcing them to rely more on the technology. The latter is surprising but may be due to the trust that has been built up by experts in the task and being familiar with the tool not catching when mistakes are made, as their very trust in the process and the tool gives them a small measure of complacency, encouraging them to rely more on the simulation (Lou & Sun, 2021).

The Theory of Technical Dominance led to a desire to measure the amount of reliance on technology. Billings et al. (1976) define two measures of overreliance on technology: Automation Bias and Automation Complacency, which several other authors tie into the Theory of Technical Dominance (Campbell et al., 2007; Parasuraman & Manzey, 2010). Automation bias is an overreliance on automation regardless of contradictory information given by a human, while automation complacency is an overabundance of trust given to the results of automation, as the system is assumed to be in a satisfactory state. Most of these papers outline that the concern is that it is usually a lack of skill or a lack of time which causes a reliance on automation to make decisions, as most of the tools are created as decision making tools. Simulations fall within this realm, as the users must assume that the designers of the simulations are built with the right assumptions and models to accurately reflect the situation that the simulation is emulating.

While technology is here to stay, it is vital that the core skills that people need are developed and encouraged (Bryant, 2011). This will be vital as it will allow the trained individual to fall back on basic skills in the cases where the technology fails. Thus, it is also important to assess, control, and monitor the technology to ensure that no entity or organization becomes entirely reliant upon it. There is an added irony where the very technology used to make our lives easier ends up making the situation worse. An example lies with sonar technology, which is used by elements within the armed forces to detect incoming airborne threats, such as missiles or enemy aircraft (Clubb, 2010). A separate study showed that two smartphones placed in the same location reported very different positions throughout the day as other devices checked on their location (Kos et al., 2013). The only measurable changes in the smartphone devices were the local weather conditions that changed throughout the day, suggesting that it is far too early to trust these devices for location finding purposes.

The failure in this example lies in the fact that the very technology that allows for a great view of the battlefield can, if relied upon as the sole means of gathering information, hinder and hurt the view of what is actually happening (Clubb, 2010). When the gaps in the technology are found or exploited, the users of that technology are often unprepared and unable to adapt quickly to escalating situations. In this case of military engagements, this can result in the loss of life. In the case of education, while not as life threatening as a physical battle would be, similar dangers can happen as students learn the use of the simulator rather than the concepts that the simulator is attempting to impart. These students, then, would be unprepared for either future courses or their future work, where other tools are more likely to be used.

A similar example occurs in automating aircraft. Dalcher (2007) describes the situation at an airshow where one of the best Air Force pilots and lead designers of a state-of-the-art aircraft attempted to demonstrate the care of the onboard safety module of the aircraft. This module was intended to stop human error from causing the aircraft to crash. Unfortunately, during the demonstration, the aircraft entered a dangerous spin after the completion of a maneuver. The very same device that should have prevented the spin from occurring also stopped the pilot from correcting the situation, as the necessary action was outside of the safe actions that the device would allow. While the pilot was able to walk away from the crash, it became apparent on further analysis that it was the box and not the pilot that caused the crash – had the pilot been able to perform the correction necessary, the plane would not have crashed. One of the strong lessons provided is that, despite the best intentions, the creation of technology to make actions safer may inadvertently do the opposite. And these scenarios are often only found through experience. They cannot be planned for.

Other areas that have been influenced by this overreliance on technology include the use of technology in travel. Now, an average consumer has access to a GPS device immediately on their mobile phone. Research has shown that people are becoming more and more reliant on this technology to navigate in their lives. There are findings that those who rely on GPS to travel are less likely to cope when the GPS fails (Grant et al., 2009). Further studies demonstrated that those who were asked to travel without GPS showed more joy in the journey and a better recollection of where they had travelled compared to those who received directions from their GPS (McCullough & Collins, 2019). This reliance on the technology appears to change the very experience that people have as they travel, making it more of a chore for those who have the technology. Javadi et al. (2017) found that the use of these navigational aids had a direct impact on the process that the brain uses to determine the paths between locations, meaning that this reliance on technology actually affects the way the brain works.

One of the other areas where there are large dangers in over-relying on technology is the medical world. Technology has had a great impact on allowing doctors and nurses to track symptoms, medication, and treatments. Often, though, there are several places where errors can arise in the process, and if one of the users falls prey to either automation bias or automation complacency, then a life could be adversely affected, potentially resulting in death (Campbell et al., 2007). The research in this area demonstrates that there are three ways that automation bias and complacency can be generated (Grissinger, 2019). First, people tend to make decisions with the least amount of cognitive effort. Second, people believe that technology is analytically superior to them. And third, people want to be able shed responsibility to the automated system. As people continue to use the same tool to perform the same task, their dependence on the tool becomes more and more engrained. The fact that they can point to the automation for errors, as well, makes them less likely to double check the results. In fact, Grissinger (2019) states that so long as the tool is correct at least 70% of the time, the use of that tool will generate a large amount of automation bias and complacency. There do not appear to be any similar examples in educational research outside of the medical field, and this is one of the gaps that this research seeks to fill.

Fortunately, the research also suggests ways to address automation bias and complacency. The designers of the tool should carefully analyze and address the potential areas of vulnerability (Bryant, 2011; Clubb, 2010; Grissinger, 2019), although this can be incredibly difficult to foresee (Dalcher, 2007). The human-computer interfaces used should be limited and designed for simplicity (Grissinger, 2019; Nahvi, 1996). The technology must be designed to reduce the over-reliance that may result from it, which may require using more than one tool for a job (Bryant, 2011; Grissinger, 2019; Lou & Sun, 2021). Training in the use of the tool must be prevalent, including what to do when the technology fails (Bryant, 2011; Clubb, 2010;

Grissinger, 2019). The technology must be designed to reduce the distractions that might occur while using the tool (Grissinger, 2019; Lou & Sun, 2021).

While automation bias and automation complacency are recognized in the literature, and many different examples have been given, most of these examples tend to be related to specific examples and situations. This makes it much harder to take the methodology and adapt it to a new environment. Definitions are shared, but the methods for measuring are varied. Deley and Dubois (2020) discuss this, demonstrating that there is a difference between trust and reliance. Trust implies that there are two separate entities and only exists if there is a possibility of betrayal of either entity to the other. Without this ability to betray, which is an active move, trust cannot exist. Technology is incapable of actively betraying trust – it follows rules, algorithms, and patterns. It does not make conscious decisions. Thus, there is no such thing as trust in technology. Instead, there is reliance, which is a continued relationship on the basis of one party's (technology's, in this case) dependable habits towards the other. A tool is made to do something. Our reliance grows as we continue to use that tool and it provides the outputs we would expect. Deley and Dubois (2020) propose using competence and predictability as measures for reliance. They point out that competence can be evaluated by the success of the tool to perform as expected, which is easy to measure. They mention that predictability is harder, partly from the vague notions behind predictability, but also because predictability is far more likely to vary from tool to tool and the uses of that tool.

An alternative way of measuring reliance, and more specifically automation bias and automation complacency, is to measure the decision-making process itself (Goddard et al., 2014). This can be done by measuring the decision accuracy, where the initial decision is measured to see if it is correct or incorrect; decision switching, the rate at which the decision is changed from the initial decision to a new decision; negative consultations, the rate at which a correct answer is converted to an incorrect answer; and pre-advice confidence, the level of confidence that the user had before receiving further advice. These measures have been shown to be effective in rating automation bias and complacency in medical workers who were given ten different scenarios, asked for an initial diagnosis, provided additional help, some of which was deliberately incorrect, and then given a chance to change their answer (Goddard et al., 2014). The focus of this research is planning to build from this model, as described below in Chapter III.

Simulations

There are three types of computer simulations (Magana et al., 2012):

1. Automated Simulations: Students watch the simulation with little/no interaction. Often used as an alternative to lectures.
2. Computer Simulations: Students use an interface to interact with the simulation. This is the most common form of simulation studies and used.
3. Computational Simulation: Students use a broad computational engine, typically one used by practicing engineers.

There are several benefits to using simulations in education. Olivas and Newstrom (1981) have shown that there are five main benefits from using simulations in a classroom setting:

1. Provide a change of pace
2. Modify the learning climate
3. Allow active participation
4. Modify the instructor role to facilitator
5. Change student attitudes

Although the focus of Olivas and Newstrom (1981) was on simulation games in social science classes, these same principles have been discussed by others within engineering education (Agnew & Shinn, 1990). Some of these other benefits include (Agnew & Shinn, 1990; Lindsay & Wankat, 2012; Uribe et al., 2016; Wankat et al., 2002):

6. Cost savings by simulating expensive equipment
7. Safer environment to perform experiments
8. Ability to make mistakes and learn from them

Based on his further research, the researcher would argue that there is another benefit to using simulations:

9. Allows students rapid self-assessment of their understanding of learning concepts

The main difference between simulations and games is that games are tools which are artificial and pedagogical, including conflict between players, rules of interactions, and goals to be achieved whereas simulations are dynamic tools, trying to represent reality, claiming fidelity, accuracy, and validity (Sauvé et al., 2007).

In their comprehensive survey of simulations and games used in higher education, Vlachopoulos and Makri (2017) argue that university instructors should be more active in the alignment of games and simulations with the curriculum while ensuring that they are implemented in blended learning environments. They should also provide scaffolding of the virtual experiences for the students. This view is further supported by Magana et al. (2012) in aligning simulations with learning outcomes. In this way, students can get the most from the Constructivist methodology and will be applying what they learn directly to the learning objectives for the course. The user interface needs to be well fleshed out and easy to navigate (Nahvi, 1996), and students should be provided with the simulation with minimal input from their instructor to really make the most of the simulation. Although students often downplay the role of the simulations in their learning, faculty can help by providing context between the simulation and the real world (Lindsay & Wankat, 2012).

It is not necessary for the simulations to be tied directly to the learning material for the course so long as the simulation is given to the students with the proper context (van der Meij & de Jong, 2006). In addition, students who had to make conscious decisions about their work rather than being given a script to follow are better able to make connections and learn from the simulations (Davidovitch et al., 2006).

Research on Simulations

Previous research studies the impact of simulations on learning (Brophy et al., 2013; Chaturvedi et al., 2011; Chyung et al., 2010; Dabbagh & Beattie, 2010; Gero et al., 2014; Goeser et al., 2011; Khan & Singh, 2015; Koh et al., 2010; Kollöffel & de Jong, 2013; Rojko et al., 2010; Tang, 2014). Many of these studies are performed as quasi-experimental studies in engineering classes with varied levels of control groups. In addition, most of them assume that their students are Caucasian male students who started college right out of high school. Some talk about how their students may not be homogenous, but few have tracked their demographics. Most articles show that students are favorably inclined towards using computer simulations. Interestingly, some studies show that older students gain more from the simulations, but that younger students prefer them more (Chyung et al., 2010). Simulation use is often followed by a measured increase in student performance as measured on quizzes and tests, but often this is simply a comparison to previous semesters rather than a true control group (Brophy et al., 2013; Goeser et al., 2011). Simulations also help keep students interested in the subject matter (Gero et al., 2014).

Simulation use inside of Civil Engineering has not been as prolific as in other engineering fields, such as Electrical Engineering. Many of the simulations involve construction management (Nikolic et al., 2011; Rokooei et al., 2017), disaster prevention (Bing et al., 2014; Fedorova et al., 2016), and component analysis of civil structures (bridges, roofs, pipelines, etc.) (Dang et al., 2017; Meschke et al., 2019; Navaee & Kang, 2017; Xiao et al., 2021). Most of these simulations are aimed at being very specific to their fields and are not introduced in early courses, such as Statics. This suggests that there is a hole in the research that could be more fully explored involving students beginning their engineering education.

Constructionism in Simulation Design

Constructionism is a core concept used in Engineering Education, where student learning occurs as students engage in the learning process (Kolari & Savander-Ranne, 2000; McHenry et al., 2005; Wankat & Oreovics, 2015). Its fundamental premise is that learning is most effective when the individual constructs his or her own understanding. Preconceptions play a large role, as they provide the foundation by which new knowledge is absorbed and understood (Kolari & Savander-Ranne, 2000). Learning should be active, continuous, and directly tied to the individual learner (McHenry et al., 2005). Learning is done by engaging students cognitively with information through activities, examples, or simulations to help them see where their understanding is flawed, incomplete, or incorrect. This is in contrast with teacher-centered modes of teaching, such as lecture, which covers the material but doesn't ensure that the students actually learn the material (Kolari & Savander-Ranne, 2000). The more engaging a concrete example is, the better the better the resulting knowledge structure will be (Wankat & Oreovics, 2015). All of this is important because it influences the way in which simulations used for education need to be designed (Nahvi, 1996). The simulation designed as the intervention for this research followed these guidelines.

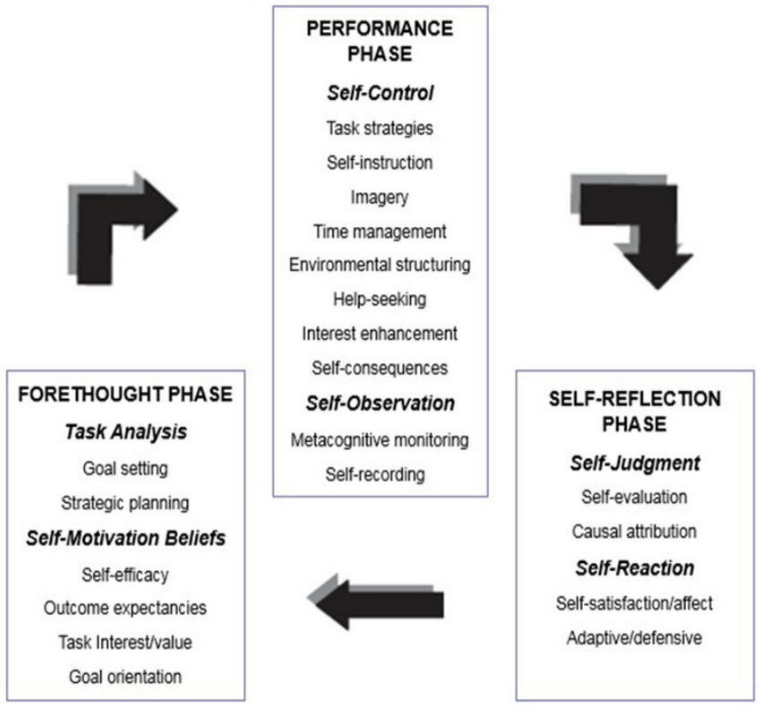
Self-Regulation in Computer Based Learning Environments

Self-regulated learning (SRL) is a conceptual framework that helps to explain the "cognitive motivational, and emotional aspects of learning" (Panadero, 2017, p. 1). That is, it attempts to tie the way that people learn and how they regulate their own learning. It was first proposed by Zimmerman (1986), and has since become a major component of educational research. In particular, Zimmerman and Moylan (2009) expanded upon Zimmerman's original theory and proposed a modified cyclical model of SRL as shown in Figure 2. Research in SRL has been explored in computer-based learning environments (CBLE) (Azevedo et al., 2016; Azevedo & Cromley, 2004; Kauffman, 2016; Narciss et al., 2007; Rivers et al., 2022), of which simulations are often a part.

As automation bias and automation complacency are directly related to the way in which a user interacts with a simulation, there appears to be a connection between SRL and automation bias and complacency. While discussing the interaction between automation of self-driving vehicles and driver's self-regulation, Wandtner (2018) argues that automation bias and complacency can lead to a lack of vigilance of the driver in monitoring the automated system. This implies that because the system is handling the task, the self-reflection or forethought phases of the cycle receive very little attention from the user and that self-regulation, then, does not occur as it should. Depending on the level of automation in a self-driving vehicle, the performance phase itself may be ignored by the driver, leading to very little actual learning taking place. Although this may be the ultimate goal of a self-driving vehicle, it does pose serious concerns for those who are using CBLEs in a learning environment.

Figure 2

Cyclical phases model. Adapted from Zimmerman and Moylan (2009).



Another concern, articulated by Azevedo et al. (2016), is whether low-achieving students can benefit from hypermedia learning environments. By utilizing a CBLE, students fail to deploy the key SRL processes and mechanisms that would have led to true learning. This is compounded by teachers who "rarely deployed scaffolding and instructional moves aimed at fostering students' self-regulated learning ecology" (Azevedo et al., 2016, p. 231).

Despite this concern, many proponents of CBLEs point out that SLR can occur in these environments. Kauffman (2016) points out that by encouraging students to self-monitor their learning, they will better engage with CBLEs. He also points out that efficacy-building feedback has a positive influence on achievement, and hopes that future research will enable more automation of this process, freeing up time for the online instructor while simultaneously encouraging more interaction from online students. Rivers et al. (2022) points out that SLR is most strongly mediated by persistence, and encourages instructors related to CBLEs to encourage students to engage with CBLEs frequently and often. Because of the change in environment, they also call for different forms of monitoring, regulation, and assessment of student performance in CBLEs. Narciss et al. (2007) showed that time spent on task in the CBLE is proportional to the student's results, showing that there is a positive impact from using CBLE. However, they were only able to evaluate the general, meta-cognitive requirements of monitoring and regulating learning activities and evaluating learning progress. They did not evaluate planning of the learning process or selecting and activating learning strategies, something they leave for future studies.

Use of Deception Studies

There are inherent dangers in using deception in studies (Boynton et al., 2015; Kemmelmeier et al., 2003; Sommers & Miller, 2013; Uz & Kemmelmeier, 2017). These dangers have been outlined above in Chapter I, and this section will show that research into simulations sometimes involves the use of deception to ensure that the study can find what it is looking for. The dangers that may arise and how they will be addressed in this research will be outlined in Chapter III below.

Uz and Kemmelmeier, (2017), point out that deception is rarely problematic so long as the deception itself is not malicious. They also stress that there are some topics that cannot be studied if the participants know the full extent of the research questions. In analyzing several studies, they also found that participants who were not in the deceived portion of a study often wish that they had been. So long as basic safeguards are in place, Uz and Kemmelmeier, (2017), argue that deception should be employed more, especially if the perspectives of the participants is taken seriously. They also stress the importance of informed consent and the opportunity for participants to reconfirm their participation in the study.

Boynton et al., (2015), found that participants who simply received false feedback, where the participant is led to believe something about themselves that isn't true or were simply deceived on the particular of the study suffered very little harm and may not generally require more than a basic debriefing. They did find that the actions of the researcher, particularly if they were of an unprofessional or insulting manner, did have a higher level of harm on the participants. This process can be overcome with proper debriefing procedures and care shown by the researcher interacting with the participants.

In terms of the debriefing itself, it is important to recognize that there are two things that must happen in the debriefing process: dehoaxing and desensitization (Sommers & Miller, 2013). Dehoaxing is the process of setting the record straight and identifying the full purpose of the study. Desensitization is the process of answering questions, alleviating stresses, and explaining the factors that are part of the study. It is vital that a good debriefing include the information that was held back, justify the use of the deception in the study, and allow the participant to withdraw when they are debriefed.

When the purpose of the study is to determine the potential overreliance on technology, it is difficult to include a study that is completely forthcoming with the participants. For example, if the participants know that they are being measured based on how critical they are of the results of a technology, they are already primed to be critical of the results. Measuring whether they would or not becomes more difficult. In a study of seventy-four studies on automation bias, at least sixteen of the included papers had some level of deception to the students (Goddard et al., 2012). In particular, the erroneous information provided by decision support systems led to an increase in negative consultations, an increase in incorrect advice cases followed, an increase in automation bias in users, and a correlation between decreasing accuracy in the system and in the user simultaneously. It is very unlikely that these studies informed users prior to the study what was being measured, as this would directly impact the reliance on the technology.

The research conducted will pursue using a deception model to be in line with these other studies and to make sure that the findings are more accurate to the participant experience. In line of the several areas that can go wrong with deception studies, this particular research will also seek to rectify the situation during the study. This will help the participants, who are students in a course learning how to perform a specific learning objective, while also potentially providing further insights into the participants' involvement and reliance of the simulation.

Summary of Literature Review

This research focused on the investigation of students' overreliance on simulations and technology. The literature review highlighted the Theory of Technology Dominance, including the supporting measures of automation bias and automation complacency, which can be used to explain the overreliance that students may have. While describing the theory, the literature also pointed out the potential dangers, ranging from a lack of information in potentially life-threatening situations to the failure to recognize when the technology has failed, and outlines potential avenues to reduce the overreliance of technology to an acceptable level of reliance. This research considered these models and advice, and it relies heavily on the concepts of automation bias and complacency to measure overreliance on technology.

Several studies have been conducted to improve STEM education by utilizing the advantages of simulations. These advantages can be leveraged to ensure that students have a good learning experience while engaging in STEM curriculum, but it is important to make sure that the dangers of using simulations is also addressed. The current research will focus on a specific simulation designed to help students learn the Method of Joints in analyzing a truss structure. This simulation was selected for many reasons. First, the concept is one that lends itself well to tool use. The concept is very process driven, enabling a computer to quickly calculate the correct answer and generate likely wrong answers. By having the simulation able to handle both correct and incorrect answers, the researcher will better be able to track the student experience and measure their reliance. Second, the simulation is being built by the researcher. The researcher has experience building simulations using a variety of programming languages, including C#, the language chosen. Although most of these simulations have been built for students in a Qualitative Reasoning course (MAT1030, which teaches students applied mathematics such as probability; expected value; dimensional analysis; algebraic modeling, including linear and exponential models; and financial mathematics, such as simple interest, compound interest, amortization, and annuities), the experience building these simulations has a direct impact on the creation of the current module. Third, the instructor of the course is closely tied in with the research, being the major advisor for the researcher. This connection will allow for a more robust simulation that can be used both to further this research but also to help students master the content presented.

A review of the literature also suggests that while metrics in overreliance in technology have been attempted in varied fields, it has not taken place within the realm of engineering education. This study would help to bridge the gap in current research between other fields, such as medicine, and engineering.

CHAPTER III

RESEARCH METHODOLOGY

The purpose of this research is to investigate the extent to which students show automation bias and automation complacency when working with a simulation. The research here will use a sequential explanatory mixed-method approach to answer the research questions. A mixed-method approach involves both quantitative and qualitative methods in a single study, either concurrently or sequentially, and are used to better understand the research being performed. For this research, we will combine a case study with quantitative analysis, allowing us to better understand differences in automation bias and automation complacency of students.

Research Questions

This research was guided by the following research questions. The third research question has two subquestions as listed below. The first two research questions were influenced by the research design of Goddard et al. (2014), which was explored in the Reliance on Technology section of Chapter 2. The researcher's expected answers to the research questions will be explained after each research question is listed here.

1. How are the participant's level of automation complacency and the correctness of the simulation that participant is using related?

Goddard et al. (2014) found that most participants switched their answer when they were highly trusting of the decision aid while simultaneously having low confidence in their own ability; similarly, the lowest number of switches occurred between the participants with high confidence in their own abilities and low trust in the decision aid. The researcher expects similar results to be shown in this study.

2. How is automation bias related to a participant's ability to recognize errors in a simulation?

Goddard et al. (2014) found that participants switched their incorrect answers to correct answers when the decision aid provided correct advice and that participants switched their correct answers to incorrect answers when the decision aid provided was incorrect. That is, most participants showed that they trusted the advice provided by the aid more than their own abilities. The researcher expects similar results in this study.

3. What factors explain the automation bias and automation complacency that the participants are experiencing?
    1. Subquestion: What factors explain the correlation between a participant's level of automation complacency and the correctness of the simulation that participant is using?
    2. Subquestion: What factors explain the impact that automation bias has on a participant's ability to recognize errors in that simulation?

This question derives itself mostly from the quantitative survey developed by Al-Natour et al. (2008) and, as outlined in the data analysis techniques found later in this section, the researcher is expecting factors to develop from responses related to trust of the simulation, seeking help during the process, being complacent with the results of the simulation, and why they switched their answers. This is outlined in Table 9, shown in the data analysis techniques section.

Researcher Positionality

It is critical that researchers recognize how their own experiences and background influence the interpretation of their results in qualitative research (Creswell & Poth, 2016). The researcher is an engineering educator, researcher, and cisgender Caucasian man, and this plays a specific role in his education and background. He has a dedicated interest in using appropriate technology in education. This interest began back in junior high school, where he was part of an experimental course that used TI-82 graphing calculators to help students learn algebra concepts. This happened in both Algebra I and Algebra II courses. Course instruction included how to use the calculators to solve the problems introduced in the content. As part of this experience, he also developed programs to help him accomplish these tasks, storing them in the calculator and making use of them throughout the course.

Part of his background also involved learning to program from a young age. As the oldest son in a family of only boys, he would spend hours watching his dad program using C++, and he taught himself how to program at a young age. His languages include Basic, qBasic, C++, C#, Java, JavaScript, HTML, and SQL. This background helped him select Electrical and Computer Engineering as his major, where he specialized in simulation work and embedded systems.

After completing a Master's degree in Electrical Engineering, the researcher took a position teaching mathematics at a local state university. This university is an open enrollment university, and the researcher taught classes primarily focused on preparing students for College Algebra. Building off of his experience, he utilized technology in his classes. This began with the Learning Management System (LMS) utilized by the university (Canvas) and the Online Homework System (OHS). The researcher has also been an advocate of teaching students to utilize their calculators, like in his own background. All of this was done with an interest of helping the students succeed.

During this time, the researcher also created additional resources to help students learn. These involved Excel files that guided students through complicated procedures, Geogebra files allowing students to explore mathematical concepts (Hohenwarter et al., 2018), and a few C# programs that also allow students to explore mathematical concepts. In the last six years, the researcher has begun teaching the Quantitative Reasoning course, focusing on providing students with mathematical content that they would see more often in the real world. The researcher began using software, such as Microsoft OneNote to aid in synchronous and asynchronous course work and Microsoft Teams to facilitate communication between students and faculty.

Through using these technologies, one thing has become clear to the researcher. Students do not always utilize the technologies correctly. Often, students implicitly trust the technology. As an example, one of the topics in the Quantitative Reasoning course is financial mathematics, where students learn about the value of compound interest (1), and amortization (2). In both cases, P represents the principal or loan amount, r represents the annual percentage rate, n represents the compound rate, and t represents the time in years. In (1), A represents the final amount of the loan due. In (2), d represents the regular deposit.

$$\mathrm{A} = \mathrm{P}\left(1 + \frac{\mathrm{r}}{\mathrm{n}}\right)^{\mathrm{nt}} \tag{1}$$

$$\mathrm{P} = \mathrm{d}\frac{1 - \left(1 + \frac{\mathrm{r}}{\mathrm{n}}\right)^{-\mathrm{nt}}}{\frac{\mathrm{r}}{\mathrm{n}}} \tag{2}$$

Because the inputs to the formulas are the nearly same, students often confuse them. In particular, many students choose to use the wrong formula. As an example, a common question is to calculate the monthly payment for a car that costs \$14,000, compounded monthly at 3.8% APR over a period of 5 years. Students often find that $P = \$14{,}000$, $r = 3.8\%$, $n = 12$, and $t = 5$. While all of these variable assignments are correct, students use (1) instead of (2) and get \$16,924.41. When asked how this represents a monthly payment for a \$14,000 car, most students reply with, “That’s what the calculator told me.” This implicit trust in the calculator belies common sense and has helped form the researcher’s interest in this topic.

It is important for a qualitative researcher to recognize that these perspectives have the potential to influence the interpretation and analysis of the results. As such, it is important to recognize the importance of using reflexivity to enhance the rigor of qualitative research (Berger, 2015; Riley, 2014). Reflexivity is the act of looking at the impact an individual has on the way they view data, taking responsibility for their own situatedness within the research, and evaluating the impact this has on the data being collected and interpreted (Berger, 2015; Riley, 2014). As such, this researcher is dedicated to ensuring that he uses reflexive practices while gathering and reviewing the data. In particular, the coding of the questionnaire and interviews will be heavily scrutinized, with the researcher going back to the collective whole of the interviews to ensure that it is the participant’s information being addressed. Further information about this process will be discussed in the section on analyzing the data.

Data Collection

Participants and Research Setting

School selection. The population of this study are students in a Statics class at Utah State University in Logan, Utah, USA. The statics class was selected for this study primarily because it is a lower-level, required class within most engineering programs, allowing for a wider variety in students. One of the instructors of this course was also the primary advisor for the researcher, providing direct oversight and understanding of the purpose of the study. The expectation is that most students will be traditional college students that match the current demographics in the engineering programs at Utah State.

Participant Selection. Students who enrolled in the Statics course at Utah State University. Participants were informed that the purpose of this study was to analyze the effectiveness of the simulation and the methods of the project related to such. They were not informed of the actual study, as the study requires that they engage in it without necessarily knowing they are being analyzed for their reliance or overreliance on technology.

The study was run over Spring 2022 and Fall 2022. Spring semester did not yield enough participants who stayed in the study after being informed of their decisions, so the study was extended to Fall semester. Between the two studies, a total of 34 participants chose to stay in the program. Instrumentation was updated between the semesters based on feedback from Spring semester. These changes are outlined in the appropriate sections below. A further advantage of running the research over two semesters is the ability to address any seasonal affects in the participants as a possible confounding variable. Demographic information about the participants can be found in Table 3.

Sample Size Analysis. G*Power was used to find the sample size that would be best for this study (Erdfelder et al., 2009; Faul et al., 2007). According to G*Power, with medium effect size, at Power=.80, and with a significance criterion at level .05, the total sample size for this study is 180. This suggests that each of the four participant group should have 45 participants in the group. The results, including an XY Plot demonstrating the Power vs Sample Size, is shown in Figure 3. Of course, having more participants in the sample would result in having higher power, as shown in the Figure 4.

The participants in the study used the simulation to work through a truss problem. Students first participated in the standard lecture provided on the topic in class. They then worked through each step of the process of analyzing a truss, with the computer providing

Table 3

Participant Demographics

Group	Subgroup	N	%
Gender	Female	9	26.5
	Male	25	73.5
Race/Ethnicity	White/Caucasian	33	97.1
	Mixed Race	1	2.9
Semester	Spring	15	44.1
	Fall	19	55.9
Academic Class	Freshmen	3	8.8
	Sophomore	25	73.5
	Junior	5	14.7
	Senior	1	2.9
Age	18-21	18	52.9
	22-24	13	38.2
	25-27	1	2.9
	31-33	1	2.9
	33+	1	2.9
Major	Mechanical	21	61.8
	Civil	9	26.5
	Environmental	2	5.9
	Electrical	1	2.9
	Biological	1	2.9
Grade Point Average	2.33-2.67	1	2.9
	2.67-3.0	2	5.9
	3.0-3.33	2	5.9
	3.33-3.67	4	11.8
	3.67-4.0	25	73.5
Prior Experience with Simulations	0-40	24	70.6
	40-80	6	17.7
	80+	4	11.8

process support (i.e., helping the participant perform the correct operation) and a check on whether the participants answer is the same as the simulation’s answer. Answers were checked at each step, and the answers were carried forward to help students in the process of solving the truss. Once the participant had completed the truss analysis, they were given an Excel file that they submitted to Canvas for further analysis. The participant population will be randomly and evenly divided into four separate groups, as outlined in Table 2.

Figure 3

*G*Power analysis*

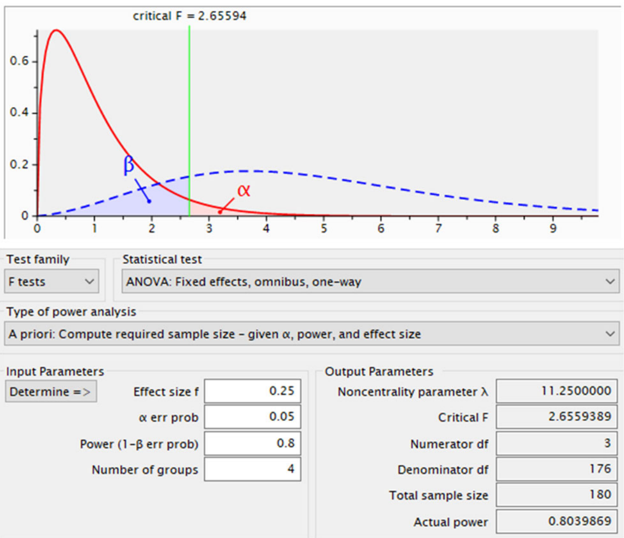
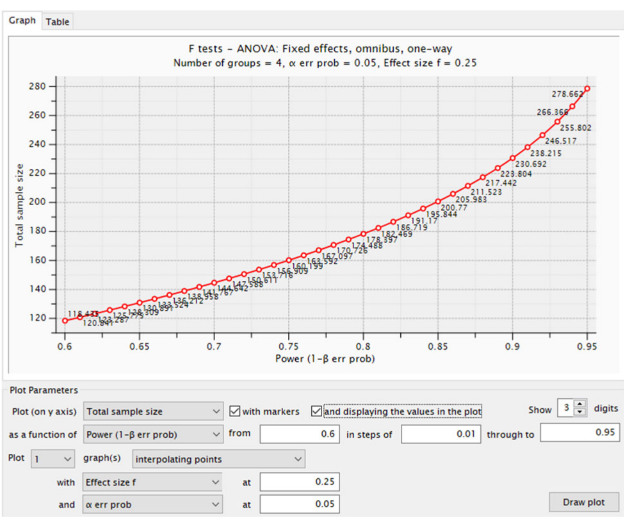


Figure 4

XY plot of power vs. sample size



The main purpose of dividing these students is to provide the framework to analyze the differences in which participants engage with, use, and rely on the simulation software. It is important to note that the two most straightforward groups are the control group, who don't know that errors could be present and won't have errors present, and misled group, who knew that errors could be present and did have errors present. The other two groups, the deceived group and the informed group, will have opposite experiences – either being deceived about whether errors were present when they are not (informed), or vice versa and not knowing errors were present when they were (deceived). Since the deceived group includes participants who were not expecting the simulation to have errors when it did have errors, these participants were the group most likely to exhibit automation bias and automation complacency, and were the group that was most likely to be negatively impacted by the deception. The informed group is the opposite – being informed of potential errors, but not actually receiving any. This group is the one least likely to exhibit automation bias or automation complacency and is intended to shed light on the ways in which students may actively work against automation bias and automation complacency. After analyzing the results of these four groups, the researcher hopes to learn more about the ways in which students rely on technology.

Instrumentations

Questionnaires have been used in many studies in educational research for several years. They provide an efficient way to gather data from a large group of people quickly (Gall et al., 2007). Questionnaires are often standardized, making the results gathered from them more objective than other methods. The researcher will use two questionnaires in this study: Demographics and Reliance on Technology questionnaires. These questionnaires are explained in detail in the following paragraphs. These questionnaires are shown in Appendices A and B.

Demographics Questionnaire. This questionnaire included: gender, age, ethnicity, year in college, and current or planned engineering program, and familiarity and experience with computer aided educational simulations. This questionnaire also included the statements on Informed Consent and some information about the study. Specifically, students will NOT be informed about the reliance side of the study, as informing them that the simulation could be intentionally wrong defeats the purpose of the study. Questions pertaining to previous simulation experience consisted of categorical data and Likert-type questions, with a single open-ended response of which software the participant has used previously. The questionnaire was built utilizing the guidance of Alchemer.com (2021), which outlines best practices in writing demographic questionnaire questions. The questionnaire was intended to take only five to ten minutes. As discussed below in the analysis section, Likert-type questions are often ordinal in nature, meaning that the data does not have clearly defined intervals between the options (Boone & Boone, 2012). As such, the central tendency of the data is best expressed as a median, the variability as frequencies, the association between values using Kendall's tau, and Chi-square tests to measure the differences between groups are significant. The researcher followed this analysis method when investigating differences associated with this questionnaire. The researcher will treat the data as ordinal during analysis. This questionnaire remained unchanged between the two semesters. The questionnaire can be seen in Appendix A.

Reliance on Technology Questionnaire. This questionnaire will inform students of the hidden aspect of the study (that errors could have been introduced), spell out which group they were in. It consists of a set of Likert scale questions adapted from Al-Natour et al. (2008). This questionnaire was used to determine how well participants benefited from a decision aid. Said decision aid helped participants purchase a laptop computer for a friend. Among other key areas, this questionnaire measures the reliance level (titled trust in the study) of the simulation. This is the most useful element of the original survey. The Cronbach's Alpha scores from the original survey for Reuse Intention (5 items), Perceived Usefulness (4 items), Trust (4 items), and Perceived Process Similarity (3 items) were .955, .951, .882, and .899, respectively (Al-Natour et al., 2008). The Domain Knowledge (4 items) element did not contain a Cronbach's Alpha but was still used as a helpful aid in understanding the results. The question responses range from 1 to 7 (*1=strongly disagree* to *7=strongly agree* or *1=very different* to *7=exactly the same*).

The researcher modified the instrument due to modify the language to make it more suited to the topic at hand. These changes were intended to change all references to the shopping assistant to a reference to the truss solving assistant, references to using the assistant to shop to using the assistant to solve truss problems, purchases with problems, and so on. Effort was made to keep the language as similar as possible, as there will not be a test of validity before participants take the questionnaire. Table 4 shows the original and modified items of the Reliance on Technology Questionnaire. The questionnaire also consists of free response questions relating to the student experience with the simulation before informing the participant of the true purpose of the study. After explaining the true purpose of the study to determine student reliance on technology, it contains questions that will help to establish the differences between the groups and to gain insights in to what they were experiencing. This questionnaire remained unchanged between the two semesters. The questionnaire can be seen in Appendix B.

Table 4

The Original and Modified Items of the Reliance on Technology Questionnaire

Category	Original items based on (Al-Natour et al., 2008)	Modified items
RI	I intend to reuse the shopping assistant for the same shopping task in the future	I intend to reuse the truss solving assistant for the same solving task in the future
RI	I predict that I will reuse the shopping assistant for the same shopping task in the future	I predict that I will reuse the truss solving assistant for the same truss solving task in the future
RI	I would consider using the shopping assistant for similar future purchases	I would consider using the truss solving assistant for similar future problems
RI	I am willing to use this shopping assistant as an aid to help with my decision about which products to buy	I am willing to use this truss solving assistant as an aid to help solve future problems

Category	Original items based on (Al-Natour et al., 2008)	Modified items
RI	I am willing to let this shopping assistant assist me in deciding which product to buy	I am willing to let this truss solving assistant assist me in solving trusses in the future
PU	Using the shopping assistant enabled me to shop more quickly	Using the truss solving assistant enabled me to solve the truss more quickly
PU	In my opinion, using the shopping assistant increased my shopping effectiveness	In my opinion, using the truss solving assistant increased my solving effectiveness
PU	In my opinion, using the shopping assistant increased my shopping efficiency	In my opinion, using the truss solving assistant increased my solving efficiency
PU	Overall, the shopping assistant was useful for shopping	Overall, the truss solving assistant was useful for solving truss problems
TR	I believe this shopping assistant is competent	I believe this truss solving assistant is competent
TR	I believe this shopping assistant is benevolent	I believe this truss solving assistant is benevolent
TR	I believe this shopping assistant has high integrity	I believe this truss solving assistant has high integrity
TR	Overall, I believe this shopping assistant is trustworthy	Overall, I believe this truss solving assistant is trustworthy
	How similar or different do you think you and the shopping assistant are in terms of:	How similar or different do you think you and the truss solving assistant are in terms of:
PS	Your decision-making style	Your problem-solving style
PS	The way you solve choice problems	The way you solve truss problems
PS	How you arrived at a decision of which a laptop is picked	How you arrived at a final answer for the truss
DK	I consider myself to be an expert on choosing computers	I consider myself to be an expert in solving truss problems
DK	I consider myself to be an expert in computer parts	-Removed due to not having a connected concept-
DK	I am knowledgeable about computers	I am knowledge about trusses
DK	I have extensive experience in buying computers	I have extensive experience in solving truss problems

Note. RI: Reuse Intention, PU: Perceived Usefulness, TR: Trust, PS: Perceived Process Similarity, DK: Domain Knowledge

It is important to note that Al-Natour et al., (2008) validated their original survey by factor analysis using Partial Least Squares and construct-item correlation using composite reliability and Fornell scores (Chin et al., 2003; Fornell & Larcker, 1981). The researcher performed this same analysis after the questionnaire to ensure that the questionnaire used was also valid.

Guided Simulation. Participants used a guided simulation that asked them to analyze joints of a truss using the Method of Joints. The original simulation was created to guide students from the beginning of a problem to the end of the problem. Participants will not be made aware that they are working through the same truss. Instead, each participant will be given an access code that will set up the program for them. The real purpose of the code is to determine which of the following solution paths the participant will work through:

1. The correct solution path: tension is treated as positive and compression as negative at the joint.
2. A solution path where the signs are reversed; tension will be treated as negative and compression will be treated as positive.

The author created the guided simulation using C#, and the process of the participant working through it will be outlined in the next subsection of the dissertation. The truss itself can be seen in Figure 10.

It is important to remember that only participants who belong to the deceived or misled groups (see Table 1) receive the code for solution path 2. Participants in the control and informed groups will all receive codes for solution path 1. Likewise, only participants in the informed and misled groups will be informed that the solution path might not be correct. This will provide the researcher with the ability to analyze the differences found in these groups, including the extent to which participants relied on the guided simulation to check their work based on what they knew about the guided simulation.

Initial Pilot/Investigation of Simulation. After the program was originally designed, a faculty member at another western university, who teaches the method of joints in an architecture class, was asked to use the program to help determine its validity. This colleague was not informed of the deceptive nature of using the program. The faculty member was given two codes, one with the correct solution path and one with the incorrect solution path. He helpfully recorded his work through the program, and met with the researcher to discuss his thoughts on the program. In particular, he pointed out that he was surprised when the program let him move on with an incorrect value, wrongly assuming that the program would not let him on until he had found a correct value. He also mentioned that the final step of checking the last joint may not check out, but the program assumes it does. He also commented that the help sections were too wordy and that they could be updated with graphics that would serve the purpose better. This feedback has led to modifications in the program to inform the participant of the usage.

Specifically, participants were informed at the beginning and then again at each step that they only have two chances to enter their values and the problem will continue without checking the value the second time. In addition, on the last joint, the program has been programmed to inform the student of whether their answers are consistent or not and will move on after one attempt, as the student cannot get a complete solution at this point. The help section has been left alone to see how if new participants feel that they would benefit from graphical rather than written help, similar to the faculty member mentioned above.

It is interesting to note that despite familiarity with solving truss problems using a method of joints, this faculty member actually made the mistake of changing the signs of their solution. They admitted that they couldn’t remember how the signs worked and so followed the help provided. They admitted to assuming that the program would need to be right and expressed surprise when informed of the true nature of the study. This aligns with Lou and Sun (2021), who mention that experts often tend to put little thought into the tasks that are assigned to the technology, as they often assume that the technology was built by another expert.

Reported data. The last step of the guided simulation will create a Comma Separated Values (CSV) file consisting of the data to be analyzed in the data analysis. This file was changed between semesters. The data gathered in Spring 2022 is shown in Table 5. The data gathered in Fall 2022 is shown in Table 6. An explanation for why these changes were necessary will be given in the next subsection on Procedures and the Modules.

Table 5

Guided Simulation Spring 2022 Data

Data name	Description
FF_{0x}	The horizontal force exerted at the fixed Joint 0, prior to checking
FF_{0y}	The vertical force exerted at the fixed Joint 0, prior to checking
FF_{3y}	The vertical force exerted at the fixed Joint 3, prior to checking
FF_{01}	The force exerted by member between joints 0 and 1, prior to checking
FF_{04}	The force exerted by member between joints 0 and 4, prior to checking
FF_{12}	The force exerted by member between joints 1 and 2, prior to checking
FF_{14}	The force exerted by member between joints 1 and 4, prior to checking
FF_{23}	The force exerted by member between joints 2 and 3, prior to checking
FF_{24}	The force exerted by member between joints 2 and 4, prior to checking
FF_{25}	The force exerted by member between joints 2 and 5, prior to checking
FF_{26}	The force exerted by member between joints 2 and 6, prior to checking
FF_{36}	The force exerted by member between joints 3 and 6, prior to checking
FF_{45}	The force exerted by member between joints 4 and 5, prior to checking
FF_{56}	The force exerted by member between joints 5 and 6, prior to checking
SF_{0x}	The horizontal force exerted at the fixed Joint 0, post checking
SF_{0y}	The vertical force exerted at the fixed Joint 0, post checking
SF_{3y}	The vertical force exerted at the fixed Joint 3, post checking
SF_{01}	The force exerted by member between joints 0 and 1, post checking
SF_{04}	The force exerted by member between joints 0 and 4, post checking
SF_{12}	The force exerted by member between joints 1 and 2, post checking
SF_{14}	The force exerted by member between joints 1 and 4, post checking
SF_{23}	The force exerted by member between joints 2 and 3, post checking
SF_{24}	The force exerted by member between joints 2 and 4, post checking
SF_{25}	The force exerted by member between joints 2 and 5, post checking
SF_{26}	The force exerted by member between joints 2 and 6, post checking

Data name	Description
SF_{36}	The force exerted by member between joints 3 and 6, post checking
SF_{45}	The force exerted by member between joints 4 and 5, post checking
SF_{56}	The force exerted by member between joints 5 and 6, post checking
RC	User confidence in their rigid force answer
J0C	User confidence in their answer to Joint 0
J1C	User confidence in their answer to Joint 1
J2C	User confidence in their answer to Joint 2
J3C	User confidence in their answer to Joint 3
J4C	User confidence in their answer to Joint 4
J5C	User confidence in their answer to Joint 5
J6C	User confidence in their answer to Joint 6
CC	User confidence in the entirety of their work

Table 6

Guided Simulation Fall 2022 Data

Data name	Description
$BEAB_1$	The easy bridge force for member AB, prior to checking
$BEAE_1$	The easy bridge force for member AE, prior to checking
$BMCD_1$	The medium bridge force for member CD, prior to checking
$BMDF_1$	The medium bridge force for member DF, prior to checking
$BHBC_1$	The hard bridge force for member BC, prior to checking
$BHBF_1$	The hard bridge force for member BF, prior to checking
$CEAB_1$	The easy cantilever force for member AB, prior to checking
$CEBD_1$	The easy cantilever force for member BD, prior to checking
$CMBE_1$	The medium cantilever force for member BE, prior to checking
$CMDE_1$	The medium cantilever force for member DE, prior to checking
$CHAB_1$	The hard cantilever force for member AB, prior to checking
$CHAD_1$	The hard cantilever force for member AD, prior to checking
$REAB_1$	The easy roof force for member AB, prior to checking
$READ_1$	The easy roof force for member AD, prior to checking
$BEAB_2$	The easy bridge force for member AB, post checking
$BEAE_2$	The easy bridge force for member AE, post checking
$BMCD_2$	The medium bridge force for member CD, post checking
$BMDF_2$	The medium bridge force for member DF, post checking
$BHBC_2$	The hard bridge force for member BC, post checking
$BHBF_2$	The hard bridge force for member BF, post checking
$CEAB_2$	The easy cantilever force for member AB, post checking
$CEBD_2$	The easy cantilever force for member BD, post checking
$CMBE_2$	The medium cantilever force for member BE, post checking
$CMDE_2$	The medium cantilever force for member DE, post checking
$CHAB_2$	The hard cantilever force for member AB, post checking
$CHAD_2$	The hard cantilever force for member AD, post checking
$REAB_2$	The easy roof force for member AB, post checking

Data name	Description
$READ_2$	The easy roof force for member AD, post checking
BEC	User confidence in their answer to Joint 0
BMC	User confidence in their answer to Joint 1
BHC	User confidence in their answer to Joint 2
CEC	User confidence in their answer to Joint 3
CMC	User confidence in their answer to Joint 4
CHC	User confidence in their answer to Joint 5
REC	User confidence in their answer to Joint 6
CC	User confidence in the entirety of their work

Interviews. The original intention was to have two participants from each group (totaling 8 participants) selected to participate in a semi-structured interview. The participants were intended to be picked in a strategic sample. The plan was to select two participants from each treatment group, with a student showing high-trust and another showing low-trust as determined by the data analysis, as outlined in the Data Analysis Techniques section below. However, as participants were asked to opt in to the interview process, only eight participants volunteered and only four actually responded to an invitation to be interviewed. This is discussed further in the findings section in Chapter 4.

The interview protocol was developed using the guidance and suggestions provided to students developing their first interview protocols (Jacob & Furgerson, 2012). The interview was created as a structured interview, with questions created ahead of time. While the interviewer was allowed to ask clarifying questions to gain further insight, all interviewees were given the same questions during the interview. This interview is intended to help shed further light on the responses from the questionnaire, with questions focused on the student reactions to the simulation, the strategies they used to work through the simulation, how they verified the results of their work, and how they feel knowing that the technology may have introduced errors (regardless of when they found out). The interview questions were face validated by a group of graduate student colleagues of the researcher. The interviews were recorded, transcribed, and coded. Coding of the interviews will use descriptive coding with *a priori* codes. The Interview Protocol can be found in Appendix C.

Procedures and the Modules

This study worked directly with human subjects, so the Utah State University Institutional Review Board reviewed the research proposal to assess the issues of risk or legal harm. This approval was received and is included in Appendix G. All questionnaires were given to participants through Qualtrics, and all questions and orders were preserved from participant to participant. In addition, both faculty members have been contacted and approved of using this study as part of their course. Informed consent was sought and received for from all students in the class wishing to participate, and an alternate assignment was provided to those students who do not wish to or could not participate in the study to relieve undue pressure that could have affected the results of the study. The activities were completed individually by the participants, and the participation occurred outside of class activities with extra credit given to students at the end of the study. The faculty members were not be informed of which participants were involved in the study nor which alternative they participated in. Scores were entered by the researcher directly in to Canvas.

Data Collection. Data collection included both qualitative and quantitative data. The primary source of data was gathered through the Demographic questionnaire and the Reliance on Technology questionnaire. The Demographic questionnaire contained quantitative results while the Reliance on Technology contained both quantitative and qualitative results. The reason for this split is to keep the quantitative benefits of previous questionnaires while also gathering qualitative information from more of the students about their experiences with the reliance portion of the experiment. Qualitative data will also be gathered from the interviews and from the results of utilizing the guided simulation.

Information about each part of the study was shared with the students via Canvas assignments and announcements. Participants will be provided with an email directly to the graduate researcher so that they can clarify questions of the research or its parts.

Questionnaires. Participants needed to complete the questionnaires online. As such, data collection was done through the use of Qualtrics questionnaires. The work did not involve providing participants with a username or a password, as the questionnaires don't require one. The results of the Qualtrics questionnaires were collected and analyzed.

Guided Simulation. The simulation provided to the Spring 2022 participants was intended to help participants work through a specific truss analysis problem while also determining the level to which the participant trusts the technology. This section will outline the simulation and how the participant interacts with it. The first screen that students saw is shown in Figure 5, including a basic explanation of the program, the notation used in the program, and a note on the notation used by the program. It was expected that students will work on their own alongside the program, which used the solution strategy presented above itself to check the participant's work. After entering the code, the participant were asked to solve for the rigid body, as demonstrated in Figure 6. The figure shows the user interface, including the rigid body of the truss, instruction text, the three data entry fields, the help and check work buttons, and the confidence radio buttons. Upon clicking the Check Work button, the user entered their confidence level and their answers will be compared to the solution path outlined above. If a different answer is present, the participant was informed of such and asked to recheck their work, including a possible suggestion to check the help. Participants only had one more chance to enter their answers. The program moved on to the next question regardless of what they entered the second time.

After the participant's confidence level was gathered and their work checked, the participant was moved on to the joints of the truss. An example of the joint analysis user interface is shown in Figure 7, including the members connected to that joint, instructions for the analyzing the joint, the data entry boxes, the help and check work buttons, and the confidence radio buttons. Each joint followed the same procedures, and the analysis was set up in such a way that there were not more than two unknown forces for any given problem. This allowed the participant the most direct path towards analyzing the entire truss. As with analyzing the rigid body, the participant recorded their confidence levels prior to having their answer compared with the current solution path. They will also only receive one additional chance to change their answer if it disagrees with the provided solution path. Each of the joints behaves in the same manner.

If a student needed help, they were first given general guidelines as shown in Figure 8. When the student requests help, the instruction text is updated with guidance towards finding a solution, including specific references to the horizontal and vertical equations needed to solve the problem and the ratios used to convert angled members to their horizontal and vertical components. All ratios (representing the angles in rise and run) are provided given the solution path outlined above; that is, participants will be told to use the appropriate sine or cosine results to find their conversion factors based on which ones the solution path wants. If a student required more detailed help, they could click on the more help button, which is shown in Figure 9. This figure contains more explanations but, more importantly, the equations used by the solution path. That is, participants will see different equations based on which solution path is tied to the code they provided at the beginning. Participants should have been able to use these equations to solve for and find the same answers as the solution path.

A completed version of the guided simulation is shown in Figure 10. Participants will be asked to measure their confidence of the entire process. The program will then tabulate the results into a .csv file that the participant will be asked to upload into the Canvas assignment.

The Fall 2022 guided simulation was changed based on feedback from the Spring participants and to bring it more in line with Goddard et al., (2014). Participants in Spring pointed out that when they made a mistake in an early part of the guided simulation, that mistake was maintained throughout the problem, leading to situations where the simulation would perpetuate other errors. It is possible that students gave up because of this error and did not even attempt the post questionnaire because of this. In addition, Goddard's et al., (2014) study had ten separate problems for their participants to engage in. As such, over the summer of 2022, the researcher redesigned the simulation to present nine different truss problems from nine different trusses, ranging in difficulty from easy 3, 4, 5 triangle trusses to medium 5, 12, 13 triangles to hard 5, 5, $5\sqrt{2}$ triangles for a bridge truss, a cantilever truss, and a roof truss. In this way, student errors on one problem were not replicated throughout other problems but each was able to be worked independently. When gathering the data shown in Chapter IV, however, it was found that

Figure 5

Welcome screen of Guided Simulation

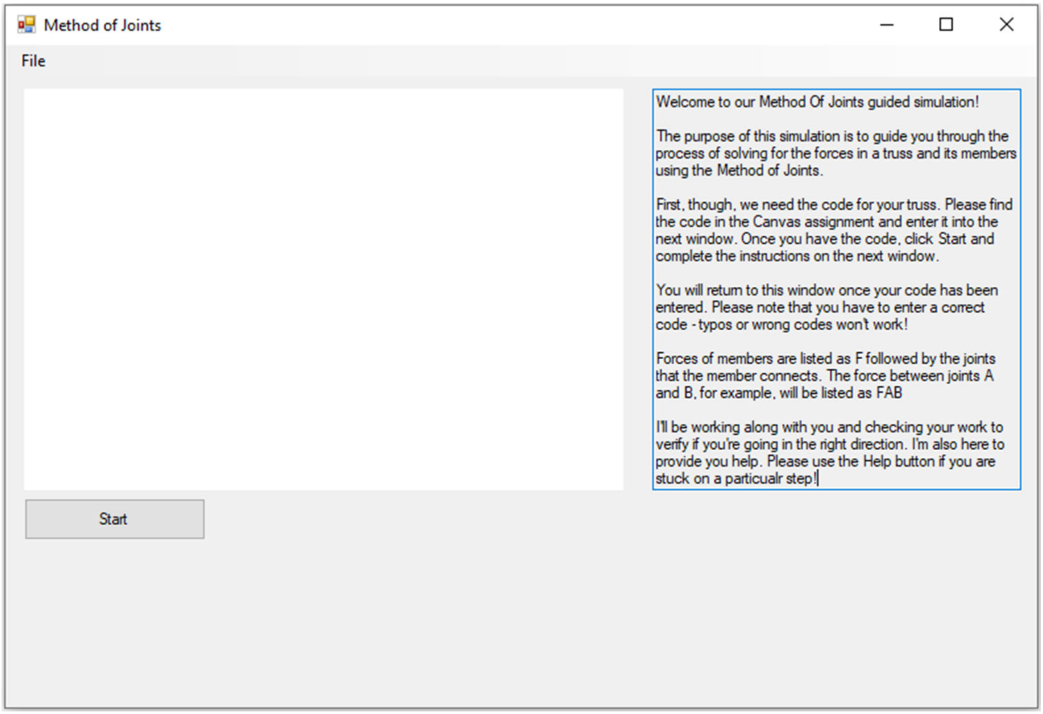


Figure 6

Solving for the Rigid Body

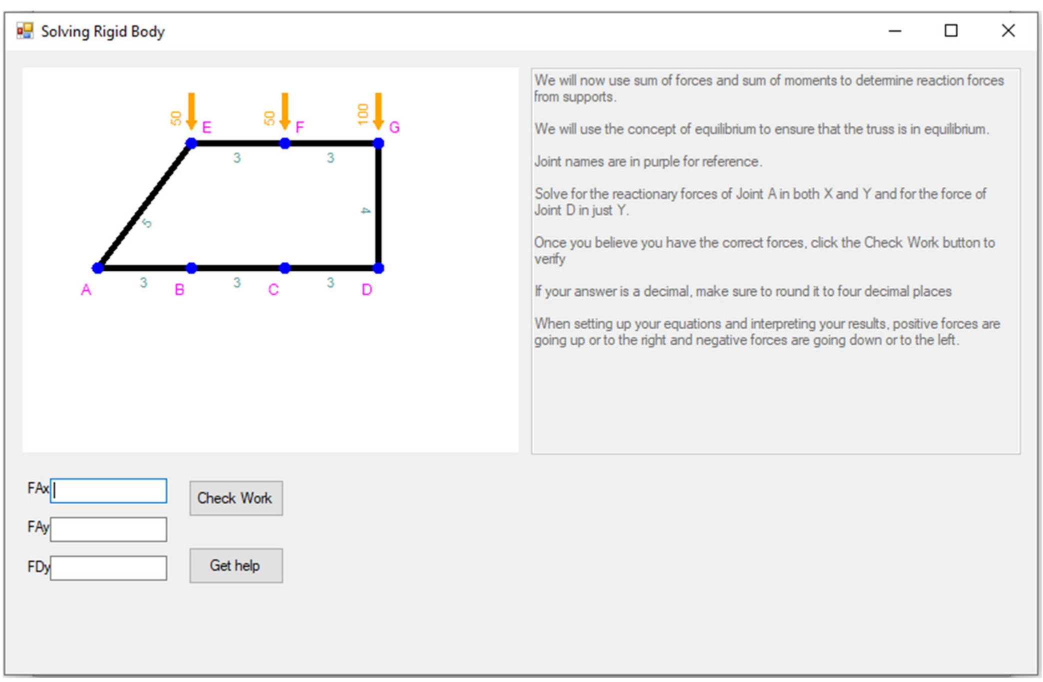


Figure 7

An Example of Solving for a Joint

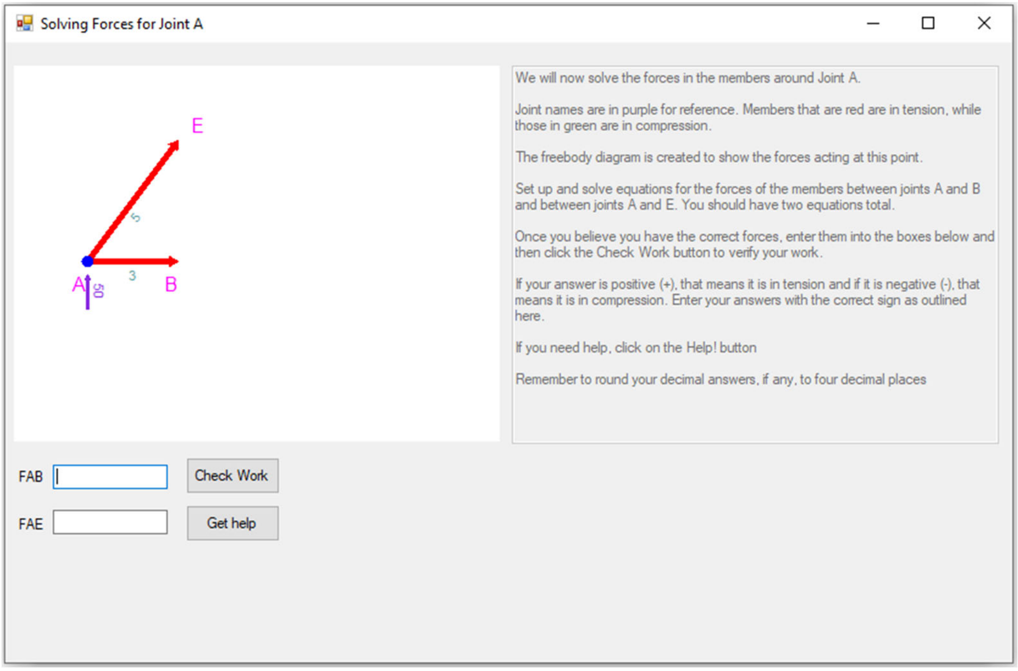


Figure 8

An Example of the Provided Help

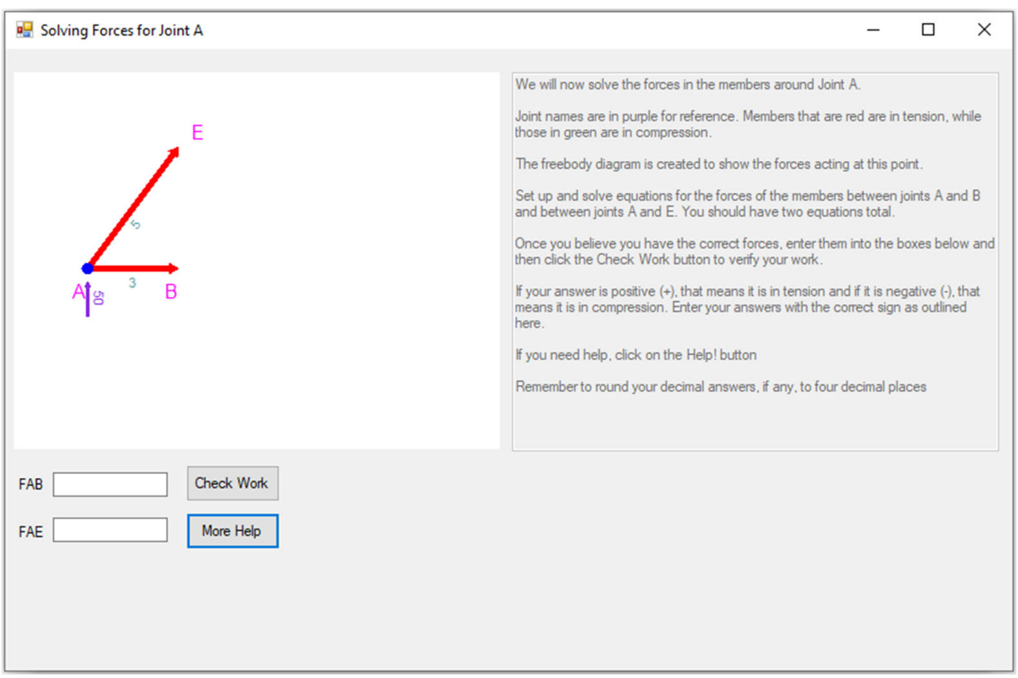


Figure 9

An Example of Receiving Advanced Help

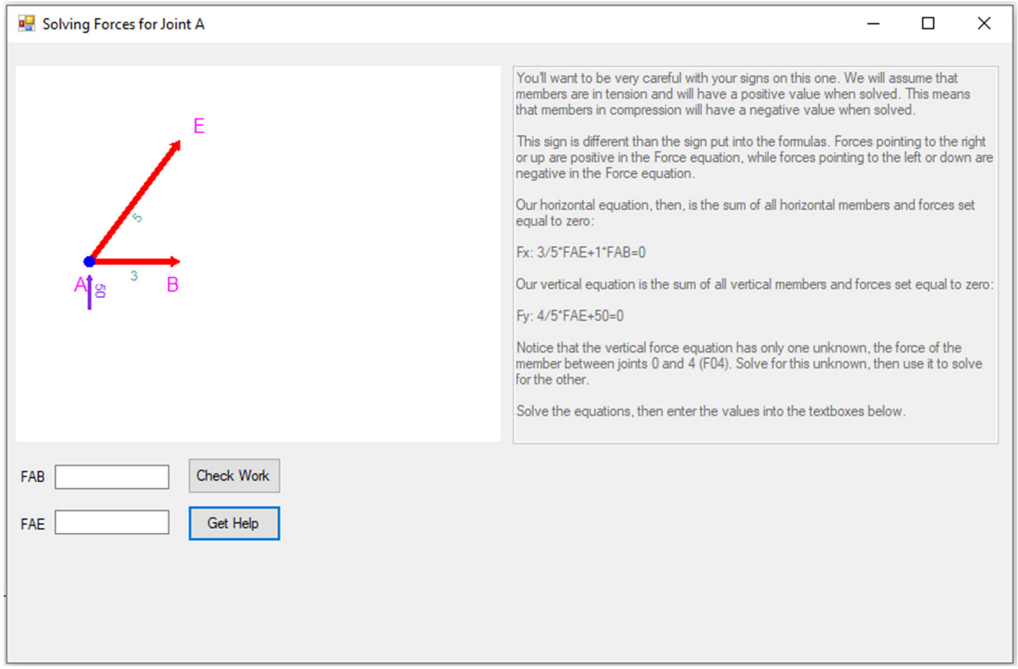
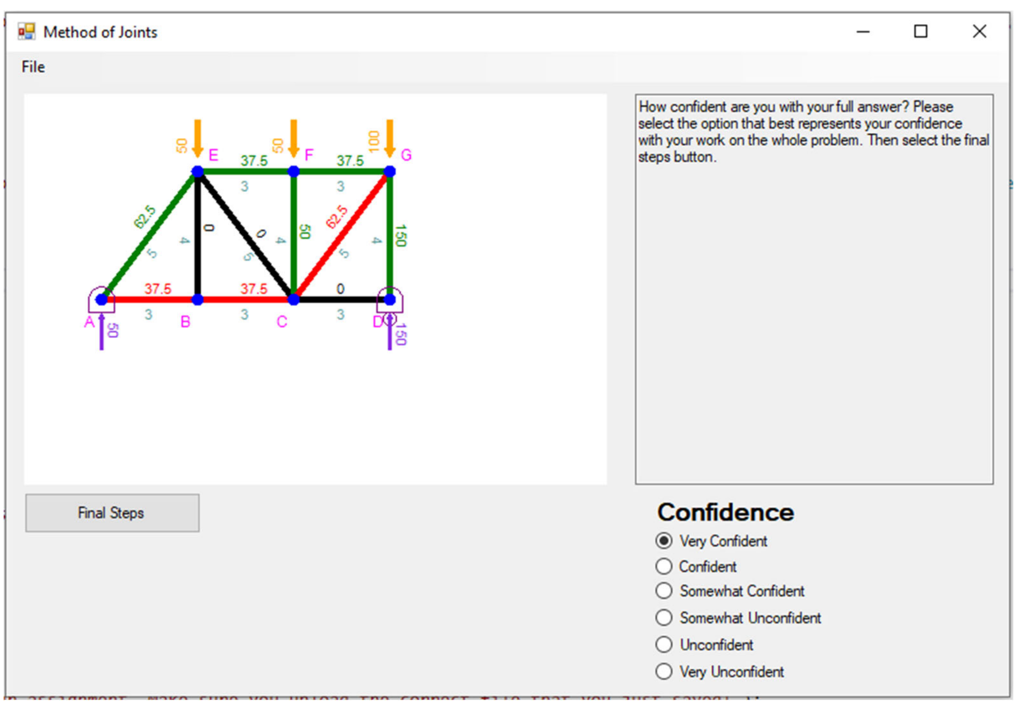


Figure 10

Example of a Completed Problem



a coding error did not adequately capture the data. In particular, the final values for the medium and hard roof truss problems were not captured, so the pre check and post check values were discarded from the data. This did allow for the same number of checks between both simulations such that the maximum number of changes made by the participants regardless of questionnaires was the same (14). The data collected is shown in Table 6. Despite updating the information shared with students in the Fall semester, three Fall participants somehow got access to Spring's guided simulation and so, for purposes of evaluation, these three students were included with the Spring Participants as opposed to the Fall participants when comparing the two simulations.

Timeline. The timeline for the research was planned to take place over a week during the mid-semester of the class. Participants began with the prequestionnaire, which should have been completed in about 10 minutes. They were then given a link to download the guided simulation and a code to access the guided simulation. Their work in the guided simulation should have taken about an hour and a half, including the upload of the results back into Canvas. The Reliance on Technology post questionnaire was expected to take a bit longer but should still have been completed within 15 minutes. For most participants, this was the end of the questionnaire, although all students who received an incorrect solution path were given the opportunity to repeat the guided simulation with the correct path to ensure that they know the proper way to attempt the problems. Those participants who were asked to participate in the interviews were contacted later in the semester. Those who agree were expected to spend another hour in the interview, although most were completed in less than half that time.

Data Analysis Techniques

After the data were collected for the four collection steps seen in Figure 1, it was reviewed to ensure that missing data were accounted for. This check focused on the questionnaires, as individual answers may be left blank. Participants who have missing values from the questionnaires were removed from the data pool as described in Chapter IV. In addition, questionnaire answers of participants was checked to ensure that none provided the same answer for each item – fortunately, no participant did so. Some of the questions were phrased with negative wording to serve as a check against this happening. The original questionnaire included some questions like this, and the updated questionnaire was done the same way. These negatively worded questions were reversed prior to data analysis. This data cleaning will be conducted prior to sampling of qualitative participants. A mapping of data sources, research questions, and analysis techniques is shown in Figure 11.

Calculating Combined Data

The results of the simulation were analyzed following the methods of Goddard et al., (2014). These values were calculated as expressed below and were used to address the research questions. These values were calculated by evaluating the results from the simulation and comparing them both against each other, and against the correct answer. One measure, the trust that the participant has of the simulation, was taken from the second questionnaire instead of from the simulation results. The specific focus on data center on the following categories, which need to be calculated:

Decision Switching. First, the number of times that the simulation results have been switched (Decision Switching) was counted (Goddard et al., 2014). To do this, the participant results were analyzed based on their correlation to the correct result both before and after being checked by the simulation. Those answers provided by the student which are correct in terms of the actual forces acting on the truss will be marked as *Right* and those that are incorrect will be marked as *Wrong*.

These designations allow for a participant's decision to keep or recalculate (switch) an answer to become classified with the following designations. A *Wrong* to *Wrong* switch (WW switch) labels a situation where the participant starts and ends with an incorrect member force, regardless of whether they changed their answer. The *Wrong* to *Right* (WR switch) switch designates where the participant starts with an incorrect member force but corrects it to a correct member force. This is noted as a positive switch. The *Right* to *Wrong* (RW switch) switch designates where the participant starts with a correct member force but corrects it to an incorrect member force. This is noted as a negative switch. The *Right* to *Right* (RR switch) switch designates a participant starting and ending with correct member forces. It is important to note that, by definition, a RR switch is not technically a switch, as each force has only one correct answer. Similarly, some WW switches may be counted as switches and others may not, depending on whether the value changed between the participant's original entry and subsequent entry after feedback from the simulation.

Each member and force exerted by the truss at a joint was assessed. It is important to note that, based on the solution path of the program, the participant was led to believe that a particular answer is right or wrong through the entire simulation, regardless of whether it actually was. This facilitates the investigation into participants trust levels for the results of the simulation and the help that they are receiving from it. A count of these switching patterns was tallied for further analysis to address the first two research questions.

The participant entered forces of each member and each support force are recorded as initial force and final force. These forces were compared against each other. Two outcomes were further explored: either both answers are the same, or they are different. If they are the same, then no switch is observed to have occurred and the count of switches for that student will not increase. If both answers are correct, then an RR switch will be recorded. If both answers are incorrect, then a WW switch will be recorded. In this way, the total number of comparisons will be accurately tracked, but decision switching won't be updated.

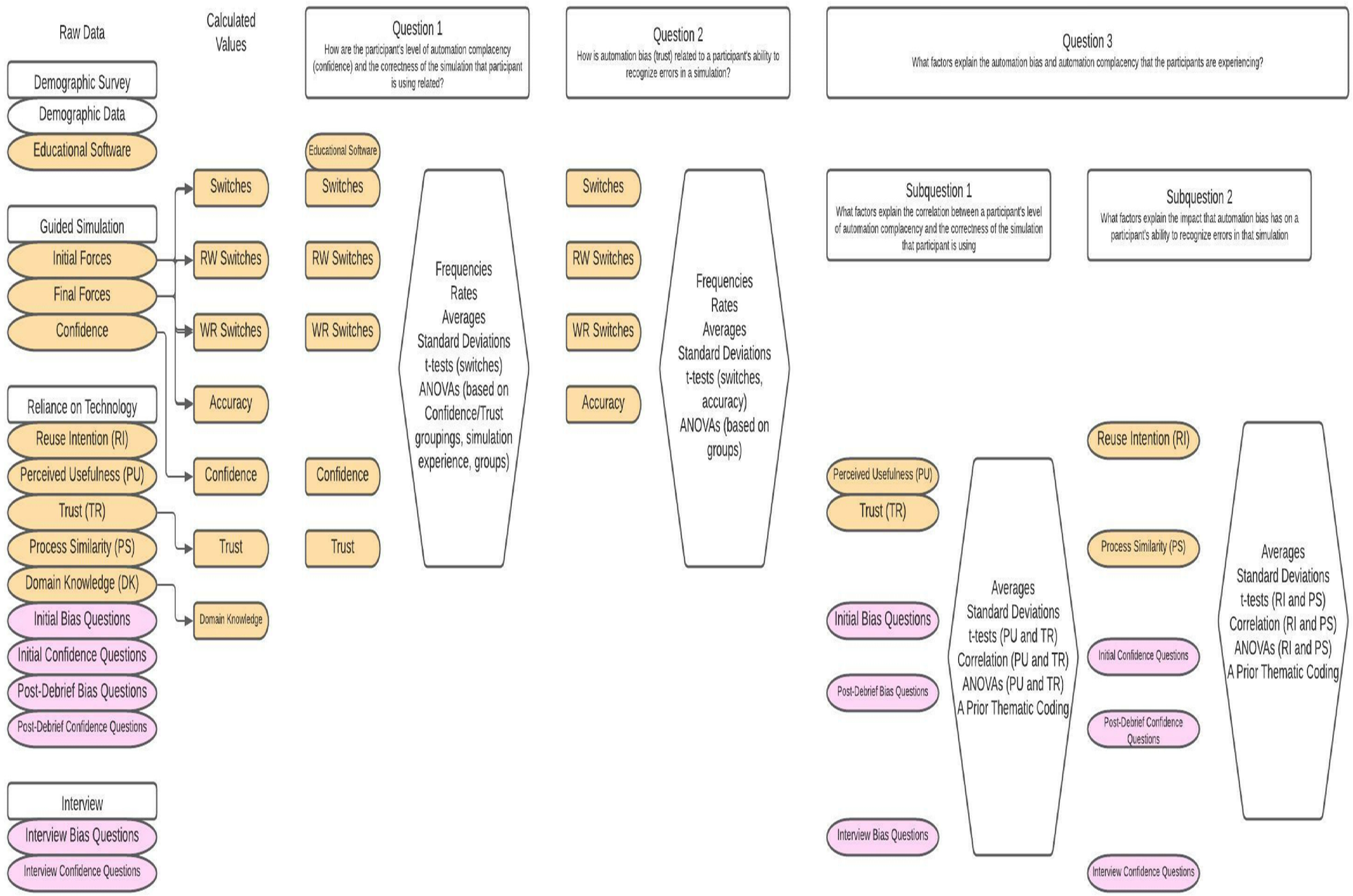
If the two values are different, then a decision switch is considered to have occurred, and the total number of decision switches will be counted. Each answer will be further analyzed to determine their classification, as outlined in Table 7. It is important to recognize that RR switches cannot occur at this stage, as the answers would both be the same – no switch has occurred. However, WW switches can occur by changing a wrong answer into another different wrong answer. A switch was considered to have occurred in this situation.

Decision Accuracy. There are two levels of accuracy that are important: whether the initial answer given to the simulation by the participant is accurate according to the actual forces exerted on the truss, and whether the final answer given to the simulation by the participant is accurate according to the actual forces exerted on the truss. The accuracy was calculated as the number of correct answers given divided by the total number of forces that should be answered for both the initial and final submissions.

Confidence. Confidence was gathered within the simulation with an imbedded six choice Likert scale question. The simulation asked for the confidence of the answers at each step in the process used to analyze the truss (i.e. after each joint is analyzed) as well as the end of the problem to understand their confidence of their work in the simulation as a whole. This work uses a six point scale taken directly from Goddard et al., (2014): Very confident-Confident-Somewhat confident-Somewhat unconfident-Unconfident-Very unconfident. This was done to keep the study close to the work of Goddard et al., (2014), allowing the researcher to validate their results. Each participant had their confidence ratings averaged into an overall rating. The

Figure 11

Research Questions Mapping



confidence of a given participant will be interpreted as high if the score ranges in the 6-4 region and low if the score ranges in the 3-1 region. This will separate the participants into highly and lowly confident groups.

Trust. Measuring the trust that the participants had of the simulation is harder to do from the simulation results. However, the fourth Trust entry from the Reliance on Technology Questionnaire specifically asked: "Overall, I believe this truss solving assistant is trustworthy." The answer to this question separated those who gave a 1-3 answer as having low trust and those with 5-7 having high trust. If a participant selected a 4 or neutral on this trust Likert-scale item, they were grouped into a neutral trust category.

Domain Knowledge. Similar to trust, this data point was generated solely from the Reliance on Technology questions on the questionnaire, which is a seven-item Likert scale question. The three domain knowledge questions from the questionnaire had their scores averaged for each student. If the average of these questions was 1-3, then the participant will be classified as having low domain knowledge about truss analysis, while 5-7 was classified as having high domain knowledge about truss analysis. If a participant selects a 4 or neutral on this Likert-scale item, the researcher will reference the accuracy of the initial force values entered by

Table 7

Mapping of Initial and Final Responses to Switch Categorization

		Initial Response	
		Right	Wrong
Final Response	Right	RR[a]	WR
	Wrong	RW	WW[b]

[a]Not a decision switch; [b]May or may not be a decision switch

the participant. Those who have an accuracy above 50% will be ranked as having high domain knowledge while those who have an accuracy below 50% will be ranked as having a low domain knowledge.

Past Experience. The demographic questionnaire has two questions about the participants' prior use of educational technology such as simulations. These data were evaluated, and those participants who showed in the top third of the data were rated as having high experience, while those who were in the bottom third will be rated as having low experience.

The ratings for Confidence, Trust, Domain Knowledge, and Past Experience will be used to help answer the research questions as outlined below. The actual statistical methods used to analyze this data will be outlined in the sections below. All of this data was calculated prior to the statistical methods being performed to answer the questions, and references to the data shown in Table 8 will have a DS prefix to draw attention to it as the data source being referenced above.

Coding of Qualitative Data

The qualitative data from the Reliance on Technology questionnaire was coded using descriptive coding to explore the topics, with an initial "a priori" list of topics related to automation complacency and bias as outlined in Table 9. These a priori codes were developed based on information found in Billings et al. (1976) and Wandtner (2018). Additional codes were looked for during the analysis, but these codes were sufficient for the research and no additional codes were found. The questions about the frequency with which participants had to recheck their work and the process will be analyzed to determine how students interacted with the simulation, and the results of those who knew there could be potential errors will be compared with those who didn't.

Table 8

Quantitative Data for Each Participant

Data Name	Explanation
DS Treatment Group	Tracks whether the participant was deceived by the program and whether they knew it. Categorized by group number (1-4).
DS Total Switches	The total number of times that the participant switched their initial answer while using the simulation.
DS RW Switches	The total number of times that a participant changed a right answer to a wrong answer.
DS WR Switches	The total number of times that a participant changed a wrong answer to a right answer.
DS Initial Accuracy	The percentage of truly correct initial answers were given
DS Final Accuracy	The percentage of truly correct final answers were given
DS Confidence	How confident the participant is in their own ability. Rated as either high (1) or low (0).
DS Trust	How trusting the participant is of the simulation. Rated as either high (1) or low (0).
DS Domain Knowledge	How knowledgeable the participant is on the topic of truss analysis. Rated as either high or low.
DS Past Experience	How much experience that the participant has with simulation software. Rated as either high or low.

Table 9

A Priori Codes for Qualitative Analysis

Automation Bias	
Trust	The participant trusted the simulation
Awareness	The participant is aware of the state of the system; i.e., if it has errors or not
Help	The participant sought help during the procedure
Automation Complacency	
Complacency	The participant checked their answers when told by the simulatio
Vigilance	The participant is actively checking to ensure the simulation is correct
Switch	The participant changed their answer

The same coding were used for the interview question data. To help ensure that the coding is done correctly, another graduate student from the Engineering Education program also participated in the coding of questionnaire and interviews. This ensured an interrater agreement that helped verify the results of the questionnaires. Eight (8) participants were selected for both coders to review, one for each simulation group from each semester. This accounted for 23.5% of the questionnaire responses being coded by both individuals. Due to the low number of interviews, all four interviews were coded by both participants. The interrater agreement is discussed in Chapter 4.

After finding an interrater agreement, the researcher continued to code the rest of the responses. Upon coding all responses, the frequencies with which the different codes appeared in answering each question were tabulated. Frequencies indicated the percentage of participants which showed what the code was intended to identify, with high frequencies showing a large number of participants demonstrating that particular code. For example, a high frequency of trust would indicate that a large number of participants express their trust in the simulation. These frequencies were analyzed to find themes related to the a priori codes and how they interacted with the participants.

Addressing Research Question #1

The first research question is, *how are the participant's level of automation complacency (confidence) and the correctness of the simulation that participant is using related?* Data calculated from the simulation results were the main measure of answering this question. This is based on the study down by (Goddard et al., 2014), who utilized decision switches to measure the complacency of medical personnel in trusting the results, even erroneous ones. To answer this question, the frequencies of the switches will be calculated for each participant. This includes the *DS Total Switches*, the number of *DS RW switches*, and the number of *DS WR switches*.

Descriptive statistics of averages and standard deviations were calculated for each *DS Treatment Group*, with a further division and analysis based on *DS Confidence*, DS T*rust*, and *DS Past Experience*. *t*-tests were performed to test for significance between the different switching groups. This was done similarly to Goddard et al., (2014), where *z*-tests were compared to determine if RW or WR switches occurred more often based on whether the right advice was given or not. This is measured through the *DS Treatment Groups*, which represents which version of the simulator the students used. *t*-tests were specifically used in this study because the researcher did not know the standard deviations of the populations (Cohen, 2008, p.203).

After calculating and testing the descriptive statistics for the participants, ANOVAs will be used to see if there are differences within and between the *DS Treatment Groups*, as well as between and within the *DS Confidence*, *DS Trust*, *DS Confidence* and *DS Trust*, and *DS Past Experience*. In particular, *DS Confidence* and *DS Trust* can both be used to show automation complacency. Someone who trusts the simulation is much more likely to be complacent with the options and will follow the advice of that simulation, while someone who is more confident in their ability is less likely to trust the simulation results which contradict that confidence. As such, while ANOVAs will be performed specifically on *DS Confidence* and *DS Trust*, ANOVAs will also be run on the crossing of *DS Confidence* and *DS Trust*. The intent was to divide participants into four additional groups for this calculation, based on being high-trust/high-confidence, high-trust/low-confidence, low-trust/high-confidence, and low-trust/low-confidence, as determined by *DS Confidence* and *DS Trust*. However, there were 31 participants who rated themselves as having high confidence, there were not enough members in the low confidence groups, so these groupings were not analyzed. As with the other analysis in this section, this is heavily influenced by Goddard et al., (2014), who found significance in trust but not confidence. This research intended to look to see if a similar pattern exists among engineering students.

Addressing Research Question #2

The second research question of this study was, *how is automation bias (trust) related to a participant's ability to recognize errors in a simulation?* As with research question #1, data from the simulation were used. The counts of *DS Total Switches*, *DS RW Switches*, and *DS WR Switches* will be used. In addition, the *DS Initial Accuracy* and *DS Final Accuracy* were analyzed and used. Descriptive statistics of averages and standard deviations of these data were calculated for each *DS Treatment Group*, which allowed us to see which groups were most likely to have automation bias. *t*-tests were used again to compare the means between groups. In addition, ANOVAs were performed to find other differences between the groups based on type of switches (*DS RW Switches* and *DS WR Switches*) and on accuracy (*DS Initial Accuracy* and *(DS Final Accuracy*).

Addressing Research Question #3

This question forms the qualitative portion while also containing some of the quantitative portion of this explanatory mixed-method design. The quantitative analysis for this question comes from the Likert-scale questionnaire results found in the Reliance on Technology questionnaire, and are intended to provide insight into the factors that will answer each question. This research question also analyzes how Self-Regulated Learning (SRL) is influenced by these same factors. Wandtner (2018) suggests that automation bias and complacency can directly influence the vigilance of the user, an inappropriate level of trust in the system, and a loss of system awareness. Each of these can contribute to a breakdown in the SRL process, particularly as there is little self-reflection or forethought done by the participant (see Figure 2Figure 1). Both a lack of vigilance and a loss of system awareness may mean that the participant is not thinking about the process well enough, losing out on appropriate forethought, while the loss of system awareness also cuts out the self-reflection phase. The participant is not evaluating the results, as the software is doing that checking for them. The codes for vigilance and awareness, shown in Table 9, are added specifically to look for these factors.

This research question has a first subquestion, *what factors explain the correlation between a participant's level of automation complacency and the correctness of the simulation that participant is using?* The results of the coding of the Reliance on Technology questionnaire and the interviews were combined with data analysis of the quantitative portion of the Reliance on Technology questionnaire to provide insights into the automation complacency that is occurring.

Specifically, based on the work of Al-Natour et al., (2008), descriptive statistics of the Perceived Usefulness (PU) and Trust (TR) questions were found and analyzed using *t*-tests between *DS Treatment Groups* on each question separately. Factor analysis was performed to ensure that the simulation performs similarly to the original work. In addition, ANOVAs were performed on PU and TR to determine if there were any significant differences in the groupings based on *DS Treatment Group*, *DS Confidence*/*DS Trust*, and *DS Past Experience*.

Finally, the qualitative coding was analyzed and used to help explain these factors, based on participant responses to the questions outlined in Table 10. The insight gained from coding each participant's responses was linked back to their answers to PU and TR to help explain the numerical results. It was also linked to the results of the analysis of Research Question 1 to help explain the results of that data. Each participant was given a unique identifier following their submission of their questionnaires and assignments so that the results can be compared with the correct participants while preserving their anonymity. In this way, the responses to the qualitative portion was used to find if the students express their thoughts the same way as they performed in the simulation and on the quantitative questionnaire questions.

The factors found by analyzing this data was also analyzed to determine what level of vigilance each *DS Treatment Group* had while using the simulation. When coupled with the other factors found, this should help provide an analysis of if the Forethought and Self-Reflection phases of the SRL were present as students worked through the simulation. This was done by looking at how often the participant responses were evaluated using the vigilance code.

Table 10

Qualitative Questions about Automation Complacency

Question	Source	A Priori Coding
Did you need to recheck your answer because the program said it disagreed with you?	Reliance on Technology, Prior to Debriefing	Complacency, Vigilance
Did you change your answer when given the opportunity to do so by the simulation? If so, why?	Reliance on Technology, Prior to Debriefing	Switch
How did you check your answer before submitting it?	Reliance on Technology, Prior to Debriefing	Vigilance
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Reliance on Technology, Prior to Debriefing	Complacency
What did you do if you disagreed with the simulation?	Reliance on Technology, Prior to Debriefing	Vigilance, Switch
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Reliance on Technology, Post Debriefing	Vigilance
How does knowing about the intentional errors impact your thoughts on the simulation?	Reliance on Technology, Post Debriefing	Complacency, Vigilance
How do you feel about having the simulation try to correct your answers now that you know it may have compared your answers with an incorrect one?	Reliance on Technology, Post Debriefing	Complacency, Vigilance

Question	Source	A Priori Coding
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Reliance on Technology, Post Debriefing	Vigilance
How would you ensure that a guided simulation or other technology you are using isn't faulty?	Reliance on Technology, Post Debriefing	Vigilance
Tell me about yourself, including your use of technology in your learning and with solving trusses using the Method of Joints.	Interview	Complacency
Tell me about your experience using the guided simulation.	Interview	Complacency, Vigilance
How often did the simulation disagree with your answer for a force when you first entered it?	Interview	Complacency, Vigilance
When the simulation disagreed with your answer, what did you do?	Interview	Switch
How did you feel when you found out the guided simulation might intentionally contain errors?	Interview	Complacency

This research question has a second subquestion, *what factors explain the impact that automation bias has on a participant's ability to recognize errors in that simulation?* To address this subquestion, the results of coding of the Reliance on Technology questionnaires and the interviews were combined with data analysis of the quantitative portion of the Reliance on Technology questionnaire.

The quantitative data gathered from the Reuse Intention (RI) and Process Similarity (PS) was analyzed using *t*-tests, just as Perceived Usefulness and Trust were for the first subquestion. ANOVAs were performed on RI and PS to see if there were any significant differences in the groupings based on their *DS Treatment Group*.

The coding performed on each qualitative question outlined in Table 11 was intended to help identify factors that help explain the automation bias of the participants that was analyzed by the second research question. As before, this data was linked to each participant by means of a unique identifier to maintain anonymity but still allowing the data to be tied to a particular individual. By combining the information gathered, the researcher hopes to be able to explain the results found in the quantitative portion.

In addition, the a prior coding of awareness was used to determine which, if any, *DS Treatment Groups* were aware of the errors in the simulation and, thus, which may have been using all three stages of SRL (see Figure 2), particularly the self-reflection phase (if they recognized the errors) and the forethought phase (if they planned how to deal with the errors). This is similar to the analysis performed in the previous subquestion in regards to vigilance.

Table 11

Qualitative Questions about Automation Bias

Question	Source	A Priori Coding
How do you feel about using guided simulations to help learn a topic?	Reliance on Technology, Prior to Debrief	Trust, Help
How did you recheck your answer before resubmitting it?	Reliance on Technology, Prior to Debrief	Trust
How often did you use the help provided by the simulation?	Reliance on Technology, Prior to Debrief	Help
Did the help provided by the simulation help you solve the truss problem? If so, how?	Reliance on Technology, Prior to Debrief	Help, Trust
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Reliance on Technology, Prior to Debriefing	Awareness
What did you do if you disagreed with the simulation?	Reliance on Technology, Prior to Debriefing	Awareness
How do you feel about the fact that the purpose of the study was not shared from the beginning?	Reliance on Technology, Post Debrief	Trust
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Reliance on Technology, Post Debrief	Awareness
How does knowing the information above about the intentional errors impact your thoughts on the simulation?	Reliance on Technology, Post Debrief	Trust
Has your perception of the help provided by the simulation changed after learning the true nature of the study and the simulation?	Reliance on Technology, Post Debrief	Help, Trust
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Reliance on Technology, Post Debrief	Trust, Awareness

Question	Source	A Priori Coding
How does this study impact your perceived usage of guided simulation moving forward?	Reliance on Technology, Post Debrief	Trust
When is the ideal time to know that a guided simulation or other technology might be faulty?	Reliance on Technology, Post Debrief	Awareness
Tell me about your experience using the guided simulation.	Interview	Trust
What was your strategy in solving the truss problem?	Interview	Trust
When the program disagreed with your answer, what did you do?	Interview	Awareness
What was your experience using the help provided by the simulation?	Interview	Help
How did you feel when you found out the guided simulation might intentionally contain errors?	Interview	Awareness

Performing the above analysis for all three research questions helped identify the confidence levels (*DS Confidence*) of the participants. The goal was to have one participant with high confidence and one participant with low confidence from each treatment group selected to participate in an interview, using the interview protocol described above. However, due to the low number of participants with low confidence and participant disinterest in participating in the questionnaire, only four participants were ultimately interviewed. Results of the interview were transcribed and coded using the same A Priori coding outlined for the Reliance on Technology questionnaire. The researcher and another graduate student from the same department then interpreted the results, triangulating the quantitative and qualitative results as shown in the tables above to establish patterns of behavior from the participants. The entire process is outlined in Figure 1.

Deception Studies

As outlined in Chapter I, deceptive studies are often used when a participant knowing what the research is about may influence the way they engage in the research. Chapter I addressed the Necessity and the Validity of the study by pointing out that a participant knowing that the study was about the reliance on technology would specifically remove the ability to see if that participant did, in fact, rely on technology. This section will discuss the other dangers of deception studies as outlined by Kemmelmeier et al. (2003). In the process of doing this, the researcher will also address how the potential harms related to the deception were handled.

Kemmelmeier et al. (2003) listed that deception can occur by commission – actively lying to participants – and by omission – lying by withholding information from participants. Both of these forms of deception occurred, although not necessarily with the same groups. This study exercised deception by omission with the control group by not informing them of the real purpose of the study. However, they were not lied to directly, as the guided simulation they worked on has the correct solution path provided to them. The deceived group also suffered from deception by omission by not knowing the purpose of the study beforehand, but in their case, they were also actively deceived by commission, as they received faulty information from the guided simulation. In particular, the simulation actively attempted to have the participants work through the problem incorrectly. The informed group had a minor deception of omission, in that the full nature of the study was not given, but they are informed that there could be faults within the guided simulation. However, there is no deception by commission, as the guided solution path was the correct one for these participants. Those in the misled group had the same minor deception of omission, but also had a deception by commission with the provided solution path in the guided simulation being incorrect. As such, all forms of deception were used in this study. The rest of this section will discuss the possible issues related to these deceptions, including ways in which the study sought to mitigate the potential harms.

Moral Turpitude

The question of moral turpitude in deceptive studies is whether the deception is morally right to use. The deontological response, based on the works of Immanuel Kant, argue that it is never correct to deceive another, while the consequentialist perspective argues that some deception is ok so long as the benefits and gains outweigh the risks. Given that the main goal of most research is the promotion of human welfare, the researcher falls more in the consequentialist perspective and believes that the deception is necessary to better understand the concept of technology dominance (American Psychological Association, 2017). As such, pressing forward with the research with the goal of making sure that the act of deception is outweighed by the gains in the research.

Violation of Participant Rights

Another question related to deceptive studies is the level to which they violate the participant's rights. That is, is the dignity of the person affected by the nature of the deception itself? Kemmelmeier et al. (2003) specifically references that one of the major reasons for Institutional Review Boards (IRBs) is to ensure that the dignity of the participants is protected, and as such, argues that deceptive studies can continue. In particular, it is important for all participants in a deceptive study to have a debriefing of the study to help them come to terms with the deception that happened.

In particular, the self-determination of the participants to participate in the study cannot be impugned by the deception itself. This can be very difficult with vulnerable groups, of which the students in the class may have fallen. Although there is little risk that those groups protected by federal law (children 18 and below, prisoners, etc.), the population may be viewed as vulnerable due to their belonging to the class and being asked to participate, with the study acting to help with a class concept. It is important that students were given the option to opt out of the study, with a similar but paper-based assignment (see Appendix F) if they so choose. In addition, participants also need to receive an update to the informed consent – with the ability to back out after discovering the true purpose of the study – and receive an apology for the deception itself. Studies that include these additions have shown that participants do not react as adversely as possible when they receive these choices, thus reinforcing the self-determination of the participants (Kemmelmeier et al., 2003). This study allowed for both options for students to self-determine how they participate in the study.

The other major concern related to the violation of the participant's rights is in the privacy of the knowledge uncovered by the deception. Because the deception is about the overreliance on technology, it did not appear that any participant's privacy will be affected by the researchers learning about the student. There was no further personal information gathered, as nothing else is germane to understanding the possible overreliance on technology. As such, the researcher sees few issues related to privacy that were discussed with the participants in the debriefing process.

Harm to Participants

There is a chance that harm may have occurred with specific participants. Participants in the deceived group ran the very specific risk that they identify the incorrect solution path as the correct solution path. This was exacerbated by the fact that it is expected that students who are new to a problem are already more likely to overly rely on the technology. Participants in the misled group may have had the same issue, but the very fact that they are informed of potential errors means that, hopefully, they were on watch for the errors and were more receptive of efforts to correct the harm. Thus, participants in the deceived and misled groups must be thoroughly debriefed on both the truth of the study and on how the program was built to deceive them on the path. This will be addressed later, when the debriefing process for this study is explained in more detail, but it is important to recognize that this is the highest degree of harm that is foreseen by this study.

In addition, it is vital to remember that there is potential for the researcher himself to be affected by participating in deception studies and that the researcher is, in some sense, a participant of the study itself (Kemmelmeier et al., 2003). The main concern expressed is that the researcher may feel emboldened to pursue further research while paying less care to the difficulties inherent to deception studies. As such, the researcher in this case reflected on the results and plans to update potential new deception studies accordingly.

Harm to Discipline

Kemmelmeier et al. (2003) argues that deceptive studies have shed light that has improved the field of psychology itself. This includes Milgram's famous study on obedience (Milgram, 1963) and research into the creation of false memory (Davis & Follette, 2001). These deception studies have greatly enhanced our understanding of psychology and are shining examples of the benefits that can be gained. The current research study will not likely have the same impact on the discipline but gaining a better understanding of the impact of overly relying on technology should have a net positive impact on the discipline.

Kemmelmeier et al. (2003) also noted that participants who participated in deceptive studies often looked at the study more critically, but there was no evidence to suggest that those same participants wouldn't take part in future studies. The researcher feels that this is the most likely outcome for those who participate in this current study, as well.

Harm to Society

Like the above arguments, Kemmelmeier et al. (2003) shows that there is no evidence of wide swaths of changes to the foundations of society using deception studies. They further point out that many experiments have benefitted from some form of deception, and that these forms of deception are praised rather than impugned. In particular, the use of blind or double-blind studies have been shown to be essential to the formation of good medical practices. The very fact that the placebo effect – the effect of positive thinking in relation to a treatment causing positive benefits – exists is further proof that society benefits from controlled levels of deception with participants. While this research will likely not reach the same impact levels that research in medical fields have, the deceptive nature of the study will help shed the light on how overreliance can occur and potentially help society to make more informed decisions related to using technology.

Risk-Benefit Assessment

The last point that Kemmelmeier et al. (2003) argues for is a detailed assessment of the potential risks, or costs, and benefits at various levels. By analyzing the costs and benefits at the levels of the participant, scientific progress, discipline, researcher, and society, it is possible to determine if the deception study will be more beneficial than harmful on the whole. It also helps to analyze those specific areas of the deception and the harm associated with it that may need repairs following the study.

In this study, the highest cost expected was at the participant level, as there are two groups that were actively being deceived. The deceived and misled groups run the risk of learning an incorrect method. While students have often learned an incorrect method before, the major risk is that the participants were actively taught an incorrect method deliberately, which may have informed their learning schemas in ways that are harder to correct later. As such, it was vital that these participants receive proper debriefing and be given the opportunity to use the guided simulation with the correct solution path given to them. On the other hand, participants in all groups may have benefited from the study, as well. They could have gained tools to help them with the specific topic but, more to the point of the study, they may also have gained an appreciation of their own level of reliance on technology. They could have potentially gained insights that may help them better approach new technologies in the future, as well.

There appears to be little potential cost to either scientific progress as a whole or to the discipline of engineering education. Similar studies have been performed which have pushed the understanding of overreliance on technology (Goddard et al., 2012, 2014; Lou & Sun, 2021). This research seeks to expand this knowledge more broadly while also specifically bringing this research to the field of engineering education.

Costs to the researcher involve the time and energy in creating a thorough research design and in interpreting the results. In addition, because it is a deception study, specific care was taken to ensure that the research minimizes the harm to the participants to the greatest extent possible. If successful, this research will open an avenue of future research using deception in Engineering Education that can help identify areas that still need additional research. Very little research has been done in Engineering Education with deception. By showing that deception research can be carefully carried out in research studies, then Engineering Education can benefit similarly to psychology, medicine, and aviation.

Finally, while there appeared to be few costs to society from this study, the benefits that could arise include more language and understanding around common knowledge, such as the importance of trying to understand the potential reliance that can arise when using a new software. In addition, educational practices could potentially be updated to take advantage of the results of this research, including ways to train new students to better understand the dangers of trusting software they don't fully understand.

Debriefing

Debriefing is the process by which participants are informed of the deception, the purpose behind the deception, and given the opportunity to react (Kemmelmeier et al., 2003; Sommers & Miller, 2013). The debriefing process for this research took place during the Reliance on Technology questionnaire following the participants' interaction with the guided simulation. This questionnaire was built so that each group received slightly different language. This questionnaire can be seen in Appendix B. The quantitative portion is the same for all participants, as are the qualitative questions. This is vital to ensure that the results of all participants can be analyzed fully. However, it is at this point that the debriefing will begin. The debriefing was modified to include information based on the group each participant was in.

Participants in the control and deceived groups were informed that there were potential errors in the guided simulation built in and of what the study was truly aiming to discover. Informed and misled groups were already made aware of this, but the purpose of the study was presented to them. The control and informed groups were informed that the solution path presented in the program was correct, while the deceived and misled groups were informed that their solution paths were incorrect. All participants were given the opportunity to continue to use the program with a correct solution path. All participants were then given further questions related to the nature of automation bias and automation complacency, as outlined in the Data Collection and Data Analysis sections above.

All participants were asked if they were comfortable continuing to be included in the study now that they knew the true nature. No penalties were given to those who choose to opt out, and their data was removed from the study according to their wishes. Participants were also given the opportunity to contact the researcher to ask any clarifying questions about the study and the deception itself, although none did. They were also asked if they are willing to participate in the interviews.

Analysis of Deceptive Study

Kemmelmeier et al. (2003) discusses the need to evaluate the costs and benefits of deceptive studies by classifying the consequences, scaling the individual outcomes, scaling the magnitudes, scaling the likelihood, and weighting the different components. This section will explore each of these evaluations.

There are far more benefits to the study than harms. The biggest harm, by far, is in relation to those participants in the deceived group, who are deceived both actively and passively. The misled group similarly has potential for harm, but it is mitigated by knowing ahead of time of the potential errors in the guided simulation. The study sought to shed light on how these participants interact with the simulation, which provided insights that could be beneficial to the discipline and to society. As such, most of the outcomes are expected to be positive.

The scale of the risks is also higher on the benefits side, as this could impact the way in which technology is used as an educational tool. By learning more about automation bias and automation complacency, it may be possible to teach students how to better prepare for using technology and how to be mindful of when they are relying on it more than they should. This will have a much smaller impact on those participating in the study, but they are expected to benefit from this, as the results of the study will be available to them to review.

Magnitude is harder to ascertain, but it certainly appears that the harm that can be caused by participants learning an incorrect method is high. The researcher ensured that these participants receive access to the correct guided simulation and an explanation about the harm that was done. As such, the researcher sought to mitigate the harm by ensuring that all participants partake in the Reliance on Technology questionnaire.

Similarly, participants in the deceived group had the highest likelihood of receiving negative consequences. They were the most likely to have an overreliance on the guided simulation, as they were the ones who had the least knowledge about the deception. As stated earlier, those new to a particular method are amongst the most likely to overly rely on the technology helping them solve the method. This further emphasizes that this group will require most reparations.

Alternative Study Methodologies

There is potential that some of this work could be found through alternative studies that do not rely on deception. However, as explained in chapter II, if the participants know the true purpose, this may guide them to make decisions differently. As such, it is unlikely that other studies will be able to provide the insights that this research is looking for. In particular, the literature review found that studies related to automation bias and automation complacency use deceptive measures inherently in their design (Al-Natour et al., 2008; Goddard et al., 2012, 2014; Lou & Sun, 2021; McCullough & Collins, 2019), where the deception occurs during the process of using the technology; or the studies have analyzed the results after the fact (Campbell et al., 2007; Clubb, 2010; Dalcher, 2007; Kos et al., 2013; Parasuraman & Manzey, 2010), where the actual thoughts of the participants during the use of the technology cannot be gathered. Only one study utilized a non-deceptive methodology, but it relied on using sophisticated functional magnetic resonance imaging (fMRI) data to analyze the results of the dependence on the technology (Javadi et al., 2017). The researcher did not have access to this level of hardware or software, nor the expertise to interpret it correctly, and so this option was not relevant. In keeping with this research, it is important to allow the deception to occur so that the participants can be measured during the process to establish more accurate findings related to their automation bias and complacency.

Summary

The use of deception in research with human subjects requires careful attention to ensure that the least amount of harm is done to participants (Boynton et al., 2015; Kemmelmeier et al., 2003; Uz & Kemmelmeier, 2017). Research on the participant's overreliance on technology requires deception to ensure that the reliance can be measured accurately. A mitigation plan is in place for all students. The deceived and misled groups, the groups that are most likely to suffer harm, will receive additional care and debriefing and will receive access to the accurate guided simulation.

The recommendations of Kemmelmeier et al., (2003), Sommers and (Miller 2013), and Uz and Kemmelmeier (2017) have been noted and followed. These include a full debriefing and the ability to withdraw from the study when the deception is revealed. The gains to the discipline are also expected to be large while the risk to the participants is small. The outcomes of the research will be evaluated to see how closely the research follows these expectations. Future research will include any remediation necessary for any deficiencies that are found.

Data Handling

Data Storage

In keeping with the Guidelines for Responsible Data Management in Scientific Research (Coulehan & Wells, 2012), all data for the study has been stored in password protected files on Box, an approved USU data storage system. The primary purpose for this was to ensure the privacy of the individual participants. By using Box, the researcher can limit who has access to what data at what times, ensuring that only those who have a valid reason to access the data have it. In particular, to preserve the anonymity of the participants, all participants in the study will be given a random identifier, and a file cross listing the two identifiers (Student A-number and assigned study identifier) was kept in a separate Box folder, also password protected. In this way, if a breach of data security were to occur, it would be harder to link individual participants with their real-life selves. The data to be protected and identified in this way includes the results of both the demographics questionnaire and the Reliance on Technology questionnaire, the report from the guided simulation, and the transcripts of those interviewed. Again, each of these items will be identified with the new study ID and not connected to the participants A-number in this area.

The interviews were conducted using Zoom, and the results were recorded and saved a separate Box folder. The files were named with the identifier created for the student and the day/time of the interview. The transcripts were kept in the same location as the other files related to that participant.

Participant's Privacy

The first step in the analysis process was to record all data in such a way that only the new identifier is on any document associated with each student. Only after the new identifier was recorded with all relevant data did the researcher analyze the data. This helped to secure the anonymity of the participants from the researcher, adding one more level of obfuscation in relation to any individual student's identity.

After the analysis began, the only time that the cross-listed document was referred to was when the students to be interviewed were identified. Once they were identified, the researcher contacted them about the interviews, conducted the interviews, and recorded the results with the assigned identifier.

Data presented in the dissertation primarily consists of amalgamated data, except in quoting specific portions of the questionnaires or interviews. No quotes will be attributed, even with the internal identifiers.

The above methods demonstrate the care used in making sure that the participant's right to privacy were protected and that the data were handled and preserved accurately.

CHAPTER IV
FINDINGS

This chapter presents the findings of the study. The findings are organized into four major sections, one for each research question with the third research question split into two separate sections, one for each subquestion. The first major section addresses how the participant's level of automation complacency (or confidence) is related to the correctness of the simulation that the participant was assigned. The second major section addresses how automation bias (or trust) is related to the participant's ability to recognize errors in the simulation. The third major section addresses the factors that can explain the automation complacency experienced by the participants, while the fourth major section addresses the factors that can explain the automation bias experienced by the participants. Additional room is given to discuss missing data, demographic information, questionnaire reliability and validity, and collected data homogeneity.

Dealing with Missing Data

As with many studies, the ways in which participants interact with the study may result in missing data. In this study, there were several places where students could opt out of the study or provide incomplete information. The first involved fully participating in the pre-questionnaire but then having incomplete participation in any further component of the study. The second place was during the post-questionnaire: students could withdraw themselves from the study after learning about the deception, forcing us to remove the data from the analysis. Finally, participants may have left questions blank or failed to provide the simulation-generated CSV file. In all these cases, students were removed from the data set because they asked (withdrawing from the study) or because they didn't provide enough information. As an example, there were blank responses provided in the second half of the Technology Questionnaire, after participants were informed of the deception and agreed to continue in the study. One participant left every question blank. One participant answered the first two questions but left the rest blank. One participant left four questions blank. One participant left three questions blank. Six participants left one question blank. These participant's data were left in the research, and their answers were ignored when coding those particular questions. With thirty-four fully completed participants, this should not have a large impact on the results of the study. Table 12 outlines the number of students who completed each questionnaire, voluntarily withdrew from the questionnaire, and those that remained, broken down by treatment group (see Table 1). The results show an almost even split by group when considering both semesters of data. Of note is the situation occurring where one of the participants in the control group completed the post questionnaire, agreed to participate in the study, but did not provide their simulation results. The researcher considered this participant as having withdrawn from the control group.

It is observable that most withdrawn questionnaires come from the groups that were given an incorrect simulation. It is also observable that more students withdrew from the second semester than the first.

Table 12

Student Engagement in Study

Group	Pre-Questionnaire	Post-Questionnaire	Withdraw	Retained
1	18 (8/10)	12 (3/9)	3 (0/3)	9 (3/6)
2	18 (8/10)	15 (5/10)	6 (1/5)	9 (4/5)
3	17 (8/9)	12 (5/7)	4 (0/4)	8 (5/3)
4	16 (7/9)	13 (6/7)	5 (3/2)	8 (3/5)
Total	69 (31/38)	52 (19/33)	18 (4/14)	34 (15/19)

Note. Numbers are presented as total students (spring students/fall students).

In addition, despite eight participants originally agreeing to be interviewed for the qualitative portion of the work, the study only resulted in four students being interviewed, one from the control group, one from the deceived group, and two from the misled group. The other four participants either did not reply to emails to schedule an appointment (three) or did not show up for either of the two times the interview had been scheduled. The results of coding the interviews will be addressed in Research Question #3 below.

Participant's Demographic Information

Analysis of the pre-questionnaire provided a description about the study participant characteristics, including gender, ethnicity, class standing, age, GPA, and major. The participants in this study numbered were thirty-four (34) (25 male and 9 female) and were enrolled at Utah State University during either the spring semester (15) or fall semester (19) in 2022. All participants self-identified as white or Caucasian, which was a single option on the questionnaire. The class is intended to be taken by sophomores, and data reflects this with 74% of participants being sophomores (24), followed by juniors (15% or 5), freshmen (9% or 3), and seniors (3% or 1). Similarly, 53% of participants were between the ages of 18 and 21 (18), 38% were 22 to 24 (13), 3% were 25 to 27 (1), 3% were 31 to 33 (1), and 3% were 33 or older (1). A majority of students (31) had a GPA of 3.0 or higher. A majority of participants declared their major as mechanical engineering (62% or 21 students), followed by civil engineering (26% or 9 students), environmental engineering (6% or 2 students), and finally electrical and biological engineering (3% or 1 student each).

Participants also self-identified their prior experience with computer aided simulations. The majority of students reported having 0 to 40 hours of experience with simulations (71% or 24), which was ranked as low experience. Student who reported 40 to 80 hours (18% or 6) were ranked as having high experience. Finally, students with 80 or more hours were ranked as having very-high experience. Students self-reported what systems they have prior experience with, and these are broken down into the following categories: Homework Systems such as Mastering Physics or ALEKS, Computational Engines such as MATLAB or Desmos, Labs such as Symbolab, LMS such as Canvas, Video hosting sites such as Khan Academy, and Programming tools such as Visual Studio. The self-reported frequencies are shown in Table 13.

It is clear from the questionnaires and the table that many students may not have understood the question on the questionnaire asking them to identify what simulation technology they had previously used. This is apparent because students are including homework systems, computational engines, labs, and learning management systems in their answers. This will be addressed in Chapter V.

Internal Reliability of Questionnaires

The internal reliability of the Reliance on Technology Questionnaire was examined using McDonald's Omega and Cronbach's Alpha (Cronbach, 1951; Dunn et al., 2014) to measure the reliability of a number of items in each category. The overall questionnaire had high McDonald's

Table 13

Computer Aided Simulations and Frequencies Reported by Students

Software	Frequency
Homework Systems	11
Computational Engines	6
Labs	4
Learning Management System	2
Computer Aided Design	3
Video	2
Programming Language	1

Note. Answers are based on student interpretation of the question about prior software use found on the prequestionnaire.

Omega and Cronbach's Alpha values, while most subgroups had good or better Omegas and Alphas. However, two subgroups, Trust and Perceived Process Similarity, had lower Omegas and Alphas. This may be a result of participants from different treatment groups having different experiences with the process and the interaction with either correct or incorrect simulation. This difference in the simulation may have led participants to answer the questions differently. The researcher expects the McDonald's Omega and Cronbach's Alpha to be lower due to the different simulations that participants used prior to completing the questionnaire. Further analysis may illustrate that the simulation groups are still homogenous, even if they differ from other groups, and this would help explain the differences found in Cronbach's Alpha. The results of testing the reliability of the questionnaire is shown in Table 14.

In addition to testing the internal reliability, the researcher further checked the face validity and the content validity of the questionnaire. The face validity was verified during the coding procedure, and the researcher determined that the answers to the questions matched the participant's groupings.

The content validity of the questionnaire was also evaluated, with the researcher and those who had oversight of the question development ensuring that all changes to the original questionnaire were those necessary to change the language from the original questionnaire

Table 14

Internal Consistency of Reliance on Technology Questionnaire (n=34)

Category	McDonald's Omega	Cronbach's alpha	*N* of items
All items	.96	.93	19
Reuse Intention	.99	.97	5
Perceived Usefulness	.95	.92	4
Trust	.63	.68	4
Process Similarity	.69	.63	3
Domain Knowledge	.87	.87	3

provided by Al-Natour et al., (2008) to the new questionnaire targeting guided simulations and the method of joints.

Data Homogeneity

Due to the importance of using homogenous data in running statistical analysis, the researcher investigated whether the data were homogenous. The process was important because the data were gathered from two different semesters with two different simulations. Tests were run to compare the different simulations used, although three students in the fall semester ended up using spring semester's simulation, as described in Chapter III, Procedures and the Modules, Guided Simulations. Data from the simulation (decision switching and accuracy) and from the reliance on technology questionnaire were used to determine if differences existed between the groups. The findings revealed that there were no significant differences based on total decision switches ($t=0.415$, $p=.681$), right to wrong switches ($t=-1.328$, $p=.200$), initial accuracy ($t=-0.036$, $p=.971$), final accuracy ($t=1.466$, $p=.152$), trust ($t=0.601$, $p=.552$), reuse intention ($t=1.268$, $p=.214$), perceived process similarity ($t=1.445$, $p=.158$), and domain knowledge ($t=0.637$, $p=.529$). However, there were significant differences found based on wrong to right decision switching ($t=2.229$, $p=.034$), where students using spring's simulation ($M=2.11$, $SD=1.97$) had significantly more changes than student using fall's simulation ($M=0.94$, $SD=1.00$), and on perceived usefulness ($t=2.056$, $p=.034$), where students using spring's simulation ($M=4.1$, $SD=0.5$) had a higher rating of perceived usefulness than fall's simulation ($M=3.8$, $SD=0.5$). With these differences in mind, the two groups are similar enough that comparisons can be made with the original groups, while consideration is given to these differences when explanations are provided.

Automation Complacency based on Simulation Correctness

This section addresses the first research question, “How are the participant’s level of automation complacency (confidence) and the correctness of the simulation that the participant is using related?” To answer the question, the researcher used the results of the computer simulation and the post-questionnaire, along with the self-reported experience with computer aided simulations. This section begins with a description of study participants through an analysis of categories based on questionnaire results.

Frequencies of Decision Switching

In order to measure automation complacency, it is vital to understand and measure whether the participant switched their answer, whether right to wrong, wrong to right, or wrong to wrong, while working in the guided simulation. This data was tabulated using Excel for all students and then grouping the switching frequencies according to simulation groups, simulation given, participant confidence rating, participant trust rating, and participant prior experience.

As outlined in Chapter III (see Table 8), the total number of switches from an initial answer were measured with particular interest given to those that were specifically Wrong to Right switches (WRSwitches) and those that were Right to Wrong switches (RWSwitches). For purposes of comparison, Wrong to Wrong switches are considered to occur whenever the initial and final answers are incorrect, regardless of whether the answer switched or not. Similarly, Right to Right switches are measures if the answer is the correct answer both before and after advice is given. Per Chapter III, neither of these latter two are used for more than reporting purposes in the tables found here. Consequentially, no t-tests or ANOVAs were performed based on wrong to wrong or right to right answer choices. Instead, complacency is best measured by determining when a correctly working simulation guides a user to change from a wrong answer to a right answer or when an incorrectly working simulation guides a user to change from a right answer to a wrong answer. The following information relates the decision switching that occurs based on the grouping frequencies outlined above.

Results for all decision switches is presented in Table 15, including those for each simulation group (control group, deceived group, misled group, informed group).

Because the simulation varied between semesters, the frequencies of decision switching was also measured based on which simulation was used. The results are shown in Table 16 for

Table 15

Direction of decision correctness before and after advice

Simulation Group	Before Advice	After Advice	Code	%	Total number
	Wrong	Wrong	WW	17.4	83
All Students	Wrong	Right	WR	11.1	53
(34)	Right	Wrong	RW	7.60	36
	Right	Right	RR	63.9	304
				100[a]	476[b]
	Wrong	Wrong	WW	19.8	25
Control	Wrong	Right	WR	9.5	12
Group (9)	Right	Wrong	RW	0.8	36
	Right	Right	RR	69.8	304
				100[a]	126[b]
	Wrong	Wrong	WW	19.8	25
Deceived	Wrong	Right	WR	9.5	12
Group (9)	Right	Wrong	RW	0.8	36
	Right	Right	RR	50	63
				100[a]	126[b]
	Wrong	Wrong	WW	10.7	12
Informed	Wrong	Right	WR	13.4	15
Group (8)	Right	Wrong	RW	0.9	1
	Right	Right	RR	75	84
				100[a]	112[b]
	Wrong	Wrong	WW	17	19
Misled	Wrong	Right	WR	11.6	13
Group (8)	Right	Wrong	RW	9.8	11
	Right	Right	RR	61.6	69
				100[a]	112[b]

[a]Total percent for group. [b]Total number of switches for group.

participants taking the simulation in the spring and fall semesters. Frequencies of decision switching were also calculated based on the self-reported confidence rating of the participants. The results are shown in Table 17. Frequencies were also calculated based on the participant's self-reported trust in the simulation. The results are shown in Table 18.

Finally, frequencies were counted based on the self-reported previous experience with simulation software. The results are shown in Table 19.

Table 16

Direction of decision switching based on semester based simulation

Semester	Before Advice	After Advice	Code	%	Total number
Spring (18)	Wrong	Wrong	WW	17.9	45
	Wrong	Right	WR	15.1	38
	Right	Wrong	RW	4.8	12
	Right	Right	RR	62.3	147
	Total			100[a]	252[b]
Fall (16)	Wrong	Wrong	WW	17	38
	Wrong	Right	WR	6.7	15
	Right	Wrong	RW	10.7	24
	Right	Right	RR	65.6	147
	Total			100[a]	224[b]

[a]Total percent for group. [b]Total number of switches for group.

Table 17

Direction of decision switching based on confidence levels

Confidence	Before Advice	After Advice	Code	%	Total number
High (31)	Wrong	Wrong	WW	16.8	73
	Wrong	Right	WR	12	52
	Right	Wrong	RW	7.6	33
	Right	Right	RR	63.6	276
	Total			100[a]	434[b]
Low (3)	Wrong	Wrong	WW	23.8	10
	Wrong	Right	WR	2.4	1
	Right	Wrong	RW	7.1	3
	Right	Right	RR	66.7	28
				100[a]	42[b]

[a]Total percent for group. [b]Total number of switches for group.

Table 18

Direction of decision switching based on self-reported trust

Trust	Before Advice	After Advice	Code	%	Total number
High (20)	Wrong	Wrong	WW	18.9	53
	Wrong	Right	WR	10.7	30
	Right	Wrong	RW	6.1	17
	Right	Right	RR	64.3	180
				100[a]	280[b]
Neutral (3)	Wrong	Wrong	WW	33.3	14
	Wrong	Right	WR	11.9	5
	Right	Wrong	RW	9.5	4
	Right	Right	RR	45.2	19
				100[a]	42[b]
Low (11)	Wrong	Wrong	WW	10.4	16
	Wrong	Right	WR	11.7	18
	Right	Wrong	RW	9.7	15
	Right	Right	RR	68.2	105
				100[a]	154[b]

[a]Total percent for group. [b]Total number of switches for group.

Table 19

Direction of decision switching based on prior experience

Experience	Before Advice	After Advice	Code	%	Total number
Very High (4)	Wrong	Wrong	WW	0	0
	Wrong	Right	WR	14.3	8
	Right	Wrong	RW	0	0
	Right	Right	RR	84.7	48
				100[a]	56[b]
High (6)	Wrong	Wrong	WW	23.8	20
	Wrong	Right	WR	8.3	7
	Right	Wrong	RW	0	0
	Right	Right	RR	67.9	57
				100[a]	84[b]
Low (24)	Wrong	Wrong	WW	18.8	63
	Wrong	Right	WR	11.3	38
	Right	Wrong	RW	10.7	36
	Right	Right	RR	59.2	199
				100[a]	336[b]

[a]Total percent for group. [b]Total number of switches for group.

Descriptive Statistics of Decision Switching

Means and standard deviations of decision switching (total switches), Wrong to Right switches, and Right to Wrong Switches, for each of the categories outlined above were calculated. The mean represents the average number of switches made by participants who meet the categorization. The results for all groupings are shown in Table 20.

t-tests and ANOVAs

In situations where only two different groupings exist within a category, such as whether the participant had a correctly or incorrectly working simulation, high or low confidence, or had the Spring or Fall simulations, t-tests were used to compare the groups. In an initial analysis, participants in the control and informed groups were combined into a correctly functioning simulation group and participants in the deceived and misled groups were combined into an incorrectly functioning simulation group. This was done to see if there were differences in switching based on the correctness of the simulation.

Table 20

Mean and standard deviations for number of switches made

Category	Subcategory	Decision Switches	Wrong to Right	Right to Wrong
Overall	(34)	5.06 (3.22)	1.56 (1.67)	1.06 (1.77)
Group	Control (9)	4.22 (3.23)	1.33 (1.58)	0.11 (0.33)
	Deceived (9)	7.00 (3.43)	1.44 (1.74)	2.56 (2.60)
	Informed (8)	3.50 (2.51)	1.88 (2.03)	0.13 (0.35)
	Misled (8)	5.38 (2.92)	1.63 (1.60)	1.38 (1.30)
Correct Simulation	Yes (17)	3.88 (2.85)	1.59 (1.77)	0.12 (0.33)
	No (17)	6.24 (3.21)	1.53 (1.62)	2.00 (2.12)
Confidence	High (31)	5.10 (3.04)	1.68 (1.70)	2.35 (2.30)
	Low (4)	4.67 (5.69)	0.33 (0.58)	3.33 (3.51)
Trust	High (20)	5.00 (3.24)	1.50 (1.67)	2.65 (2.32)
	Neutral (3)	7.67 (3.51)	1.67 (1.53)	4.67 (3.51)
	Low (11)	4.45 (3.05)	1.64 (1.86)	1.45 (1.81)
Experience	Very High (4)	2.00 (2.83)	2.00 (2.83)	0 (0)
	High (6)	4.50 (2.81)	1.17 (1.17)	3.33 (2.50)
	Low (24)	5.71 (3.16)	1.58 (1.61)	2.63 (2.34)

Note. Subcategory numbers represent number of participants in each group.

The first set of t-tests was performed to compare the means of groups separated into who used each of the two simulations, either the one in Spring semester or the one in Fall semester. The results are shown in Table 21.

The key finding of these t-tests is that only the number of Wrong to Right switches were statistically significant, with students making more Wrong to Right switches in Spring's simulation than were made in Fall's simulation (p<.05). Spring's simulation involved participants solving an entire truss while Fall's simulation involved participants solving nine separate joints from nine different trusses.

The second set of t-tests was run to determine if there were differences in the mean number of switches between students with high confidence and students with low confidence. These tests were run on the total number of decision switches, the number of Right to Wrong decision switches, and the number of Wrong to Right decision switches. The results are shown in

Table 21

t-test results for decision switching

Grouping	Decision Switch Category	Label	n	t-statistic	df	p
Semester Simulation	Total Switches	Spring, Fall	18, 16	0.42	32	.681
	RW Switches[a]	Spring, Fall	18, 16	-1.33	19.54	.200
	WR Switches[a]	Spring, Fall	18, 16	2.23	25.82	**.035***
Confidence	Total Switches	High, Low	31, 3	0.22	32	.829
	RW Switches	High, Low	31, 3	0.06	32	.953
	WR Switches	High, Low	31, 3	1.34	32	.188
Correct simulation	Total Switches	Yes, No	17, 17	-2.26	32	**.031***
	RW Switches[a]	Yes, No	17, 17	-3.61	16.78	**.002****
	WR Switches	Yes, No	17, 17	0.10	32	.920

Note. Subcategory numbers represent number of participants in each group

[a]Levene's test for Homogeneity of Variance was violated; since the variances are not the same, these tests were performed using Welch's Two Sample t-test.

**p*<.05. ***p*<.01.

Table 21. The table shows that there were no differences between groups based on confidence. It should be pointed out that there is a low number of low confidence participants, and this may have had an impact on the results.

The final set of t-tests was run by comparing the number of switches of all participants in the correctly working simulation against those in the incorrectly working simulation. This was done to see if differences existed in the mean number of decision switches based on the presentation of the simulation to the respective groups, i.e., whether it gave correct advice or incorrect advice. The results of the t-tests are shown in Table 21.

The table shows that participants who had an incorrect simulation switched any of their answers and, more specifically, from right answers to wrong answers statistically significantly more often than those who had a correct simulation.

Because many of the groupings under consideration had more than two comparative levels, ANOVAs were performed to determine if there were differences in the means between the groups and, if there were, where they occurred. These ANOVAs were done to check for significance between the treatment groups, self-declared trust levels, and self-declared experience.

The first ANOVA was performed based on which treatment group the participant belonged to, whether the control, deceived, informed, or misled groups (see Table 1). The results of the ANOVA are shown in Table 22.

The ANOVA found that there were significant differences in the number of Right to Wrong answer switches but not in the number of Total Switches or in the Wrong to Right switches. Post hoc tests were performed using Tukey's HSD test, which showed that treatment control group had a significantly lower mean than the deceived group (p=.009) and the informed

Table 22

ANOVA results for decision switching

Grouping	Decision Switch Category	df	F	η	p
Treatment Group	Total Switches	3, 30	2.15	.177	.115
	RW Switches[a]	3, 30	5.31	.347	**.005***
	WR Switches	3, 30	0.16	.015	.926
Trust	Total Switches	2, 31	1.20	.072	.316
	RW Switches	2, 31	0.32	.020	.726
	WR Switches	2, 31	0.03	.002	.972
Prior Experience	Total Switches	2, 31	2.62	.145	.089
	RW Switches[a]	2, 31	2.80	.153	.076
	WR Switches	2, 31	0.29	.019	.748

[a] Levene's test for Homogeneity of Variance was violated; Welch's test confirms the significance despite this violation.

*$p<.05$.

group had a significantly lower mean than the deceived group (p=.012). More participants in the deceived group changed a correct answer to an incorrect answer once feedback was provided by their incorrect solution.

The second ANOVA was performed based on the trust, whether that is high, neutral, or low, that participants self-declared about the guided simulation. The results are shown in Table 22. These results show that there is no difference in switching based on how much trust is placed on the technology by the student.

The final ANOVA was performed comparing participants' decision switching by comparing them against their self-reported prior experience with simulations. The results are shown in Table 22. The ANOVA shows that there are no statistically significant differences based on prior experience.

Automation Bias and Participant Perception

This section addresses the second research question, "How is automation bias (trust) related to a participant's ability to recognize errors in a simulation?" To answer this question, the researcher used the results of the computer simulation, in particular the number of decision switches, Right to Wrong switches, Wrong to Right switches, and the initial and final accuracy of participants' answers to each question. Because the number of decision switches, Right to Wrong switches, and Wrong to Right switches have already been calculated to answer the first question, this information will not be presented here. Instead, the results from above will be coupled with the analysis of the participant's initial and final accuracy as based on their responses to each part of the guided simulation. This section begins with a description of study participants through an analysis of the categories based on questionnaire results.

Descriptive Statistics of Answer Accuracy

In order to measure automation bias, it is important to measure both whether decision switching occurred while working in the program and whether the initial and final forces provided by the participant are correct. Automation bias measures the extent that the participant trusts the technology more than their own ability. We would expect those exhibiting automation bias to have their final answers match with the correctness of the simulation they are using – those with a correctly working simulation should show higher final accuracy than those with an incorrectly working simulation. This data was calculated using R for all participants, then calculated for each group according to simulation groups, the simulation given, participant confidence rating, participant trust rating, and the participant's prior experience using simulations.

As outline in Chapter III (see Table 8), the initial and final accuracy of the participant's answer for each force they were solving for in the truss was measured. Accuracy is based on the participant correctly solving for the provided forces. Both accuracies are calculated by counting the number of correct responses and dividing by the number of forces submitted (14). This provides an overall accuracy for the participant for both their initial and final submissions. The initial accuracy comes from measuring the correct number of responses given by the participant before receiving feedback from the simulation while the final accuracy comes from measuring the correct number of responses after receiving feedback from the simulation. These accuracies are then compared to determine if there are differences between participants based on the groups. In addition, the difference between the final and initial accuracy was found to show the change in accuracy resulting from the feedback from the simulation. The means and standard deviations of these accuracies are presented for all groups in Table 23.

Table 23

Descriptive statistics for initial and final accuracies

Grouping	Subcategory	Initial Accuracy (%)	Final Accuracy (%)	Change (%)
Overall	(34)	56.1 (23.6)	59.7 (28.3)	3.6 (18.0)
Group	Control (9)	49.2 (28.9)	57.9 (30.0)	8.7 (11.7)
	Deceived (9)	50.0 (24.2)	42.1 (23.5)	-7.9 (21.3)
	Informed (8)	68.8 (20.9)	81.2 (20.2)	12.5 (15.2)
	Misled (8)	58.0 (16.4)	59.8 (28.0)	1.8 (17.8)
Correct Simulation	Yes (17)	58.4 (26.6)	68.9 (27.8)	10.5 (13.2)
	No (17)	53.8 (20.7)	50.4 (26.5)	-3.4 (19.7)
Confidence	High (31)	58.8 (22.0)	63.1 (26.2)	4.4 (18.5)
	Low (4)	28.6 (25.8)	23.8 (28.9)	-4.8 (8.2)
Trust	High (20)	55.7 (25.4)	60.4 (28.4)	4.6 (15.2)
	Neutral (3)	38.1 (36.0)	40.5 (47.6)	4.8 (18.0)
	Low (11)	61.7 (14.7)	63.6 (23.1)	1.9 (23.5)
Experience	Very High (4)	67.9 (13.7)	82.1 (18.0)	14.3 (20.2)
	High (6)	52.4 (30.5)	60.7 (28.1)	11.9 (8.4)
	Low (24)	55.1 (23.3)	55.7 (28.9)	3.9 (18.9)

Note. Subcategory numbers represent number of participants in each group.

t-tests and ANOVAs

Where only two groups exist within each category, such as whether the participant had a correctly or incorrectly working simulation, high or low confidence, or had the Spring or Fall simulations, t-tests were used to compare the groups. As with research question one, above, the control and informed groups were combined into a correct simulation group and the deceived and misled groups were combined into a correct simulation group.

The first t-test was performed based on whether the participant was using the Spring simulation or the Fall simulation. The results are shown in Table 24. This shows that there were no significant differences found based on the simulation used for either the initial accuracy or the final accuracy.

The second t-test was performed based on the confidence rating of the participants. The results are shown in Table 24. This shows that there is a significant difference favoring those with high confidence in their abilities and those with low confidence in their abilities, with those with high confidence having higher initial and final accuracies, where accuracy is the number of correct responses divided by the number of forces solved, than those with low confidence.

The final t-test performed was based on whether the participant was placed into a correct simulation group or not, as defined at the beginning of this subsection. The results are shown in

Table 24

t-test based on Spring or Fall simulation and accuracy

Grouping	Accuracy Category	Simulation	n	t-statistic	df	p
Semester	Initial Accuracy	Spring, Fall	18, 16	-0.04	32	.971
Simulation	Final Accuracy	Spring, Fall	18, 16	1.47	25.82	.152
Confidence	Initial Accuracy	High, Low	31, 3	2.24	32	**.032***
	Final Accuracy	High, Low	31, 3	2.46	32	**.019***
Correct	Initial Accuracy	Yes, No	17, 17	0.57	32	.576
Simulation	Final Accuracy	Yes, No	17, 17	1.98	32	.056

$^{*}p<.05$

Table 24. The results show no significant differences in initial accuracy based on whether the students had a correctly working simulation or not, although final accuracy appears to be close to significance, but didn't hit the expected threshold (*p* was between .1 and .05).

ANOVAs were performed for any grouping with more than two groups. These ANOVAs were done to check for significance between the treatment groups, self-declared trust levels, and self-declared experience. The first ANOVA was performed based on which treatment group the participant was in. These ANOVAs were done to check for significance between the treatment groups, self-declared trust levels, and self-declared experience. The results are shown in Table 25.

The ANOVA showed that there was a significant difference between the control, deceived, informed, and misled groups. Tukey's HSD was used to perform Post Hoc tests on the groups and found a significant difference in final accuracy between the deceived and informed groups (p=.019), with those in deceived group having lower final accuracy than those in the informed group. No other group comparisons were significant.

The next ANOVA performed was based on the participants' self-declared trust ratings. The results are shown in Table 25. The results show that there were no significant differences for initial accuracy or final accuracy based on the participants' trust ratings.

Table 25

ANOVA results based on accuracy

Grouping	Accuracy Category	df	F	η	p
Treatment Group	Initial Accuracy	3, 30	1.27	.113	.302
	Final Accuracy	3, 30	3.28	.247	**.035***
Trust	Initial Accuracy	2, 31	1.20	.072	.315
	Final Accuracy	2, 31	0.79	.049	.462
Prior Experience	Initial Accuracy	2, 31	0.58	.036	.566
	Final Accuracy	2, 31	1.55	.091	.228

*p<.05

The final ANOVA performed was based on participant prior experience. The results are shown in Table 25. The results show that there were no significant differences based on prior experience.

Factors Leading to Automation Complacency

This and the next section address the third research question, which is "What factors explain the automation bias and automation complacency that the participants are experiencing?" This section will specifically address the subquestion related to automation complacency, which is "What factors explain the correlation between a participant's level of automation complacency and the correctness of the simulation that participant is using?"

Descriptive Statistics of Perceived Usefulness and Trust

Means and standard deviations for each subquestion related to Perceived Usefulness and Trust in the Reliance on Technology questionnaire were calculated and are presented below in the table in this form, mean (SD). To help analyze the results, the average Perceived Usefulness

Table 26

Descriptive Statistics for Perceived Usefulness and Trust Overall

Category	Subcategory	Perceived Usefulness	Trust
Overall	(34)	4.1 (1.6)	4.3 (1.2)
Group	Control (9)	5.2 (1.2)	5.2 (0.9)
	Deceived (9)	3.8 (1.5)	3.6 (1.4)
	Informed (8)	4.6 (1.6)	4.8 (0.8)
	Misled (8)	2.7 (1.2)	3.5 (0.9)
Correct Simulation	Yes (17)	4.9 (1.4)	5.0 (0.9)
	No (17)	3.3 (1.5)	3.6 (1.1)
Confidence	High (31)	4.1 (1.7)	4.3 (1.3)
	Low (4)	4.2 (0.8)	4.1 (0.9)
Trust	High (20)	4.7 (1.4)	5.0 (0.8)
	Neutral (3)	3.9 (1.0)	4.2 (0.8)
	Low (11)	3.0 (1.7)	2.9 (0.8)
Experience	Very High (4)	4.8 (1.3)	3.9 (0.7)
	High (6)	4.5 (1.8)	4.6 (1.2)
	Low (24)	3.9 (1.7)	4.3 (1.3)

Note. Subcategory numbers represent number of participants in each group.

and Trust was found for each participant as designated by constructs in the questionnaire. The results for both averaged results are found in Table 26, while individual question results are shown in Table 27 for individual questions for the constructs of Perceived Usefulness and Table 28 for Trust.

t-tests and ANOVAs

As with both questions before, where only two categories existed within a grouping, as is the case with the correctness or incorrectness of the simulation and the self-declared confidence level of the participant, t-tests were used to look for differences while ANOVAs were used when three or more categories existed. Similarly, participants in the control and informed groups were combined into a single, correctly working simulation while students in the deceived and misled groups were combined into a single, incorrectly working simulation.

Table 27

Descriptive Statistics for Perceived Usefulness Questions

Category	Subcategory	Solve Quickly	Decrease Effectiveness[a]	Decrease Efficiency[a]	Useful Overall
Overall	(34)	3.9 (1.9)	4.0 (1.8)	4.3 (1.8)	4.1 (1.7)
Group	Control (9)	5.2 (1.4)	4.9 (1.7)	5.4 (1.0)	5.2 (1.2)
	Deceived (9)	3.7 (1.8)	4.0 (1.8)	4.1 (1.8)	3.6 (1.9)
	Informed (8)	4.0 (2.3)	4.6 (1.8)	5.0 (1.9)	4.8 (1.3)
	Misled (8)	2.8 (1.7)	2.5 (1.2)	2.6 (1.4)	2.9 (1.5)
Correct Simulation	Yes (17)	4.6 (1.9)	4.8 (1.7)	5.2 (1.4)	5.0 (1.2)
	No (17)	3.2 (1.8)	3.3 (1.7)	3.4 (1.8)	3.2 (1.7)
Confidence	High (31)	3.9 (2.0)	3.9 (1.9)	4.4 (1.9)	4.2 (1.7)
	Low (4)	4.3 (1.2)	5.0 (1.0)	4.0 (1.0)	3.3 (1.5)
Trust	High (20)	4.5 (2.0)	4.6 (1.6)	5.1 (1.3)	4.6 (1.2)
	Neutral (3)	3.0 (1.0)	3.0 (1.0)	4.3 (2.1)	5.3 (0.6)
	Low (11)	3.1 (1.8)	3.2 (2.0)	2.9 (1.9)	2.9 (2.0)
Experience	Very High (4)	4.2 (2.3)	4.5 (1.8)	5.2 (1.6)	4.2 (1.7)
	High (6)	3.9 (1.9)	3.6 (1.8)	4.0 (1.9)	4.0 (1.7)
	Low (24)	4.0 (1.8)	5.8 (1.3)	5.0 (1.8)	4.5 (2.1)

[a]Reversed question in original questionnaire – results switched to treat high values as good.

Table 28

Descriptive Statistics for Trust Questions

Category	Subcategory	Competent	Benevolent	Low Integrity[a]	Trustworthy
Overall	(34)	3.9 (1.9)	4.5 (1.6)	4.5 (1.6)	4.3 (1.8)
Group	Control (9)	5.0 (1.7)	4.9 (1.8)	5.7 (0.9)	5.1 (1.7)
	Deceived (9)	3.4 (2.0)	4.1 (1.9)	3.6 (1.9)	3.4 (2.0)
	Informed (8)	4.6 (1.8)	4.2 (1.2)	5.1 (0.8)	5.4 (0.7)
	Misled (8)	2.5 (1.4)	4.6 (1.3)	3.6 (1.4)	3.1 (1.6)
Correct Simulation	Yes (17)	4.8 (1.7)	4.6 (1.5)	5.4 (0.9)	5.2 (1.3)
	No (17)	3.0 (1.8)	4.4 (1.6)	3.6 (1.7)	3.3 (1.8)
Confidence	High (31)	3.9 (2.0)	4.5 (1.6)	4.5 (1.6)	4.3 (1.9)
	Low (4)	3.7 (1.5)	4.7 (1.2)	4.0 (1.7)	4.0 (1.7)
Trust	High (20)	5.0 (1.5)	4.5 (1.5)	5.0 (1.3)	5.6 (0.7)
	Neutral (3)	3.7 (1.2)	4.3 (1.2)	5.0 (1.0)	4.0 (0.0)
	Low (11)	2.1 (1.5)	4.4 (1.9)	3.4 (1.7)	1.9 (0.7)
Experience	Very High (4)	2.2 (1.0)	5.2 (1.5)	4.8 (1.9)	3.5 (2.4)
	High (6)	4.5 (2.4)	4.5 (1.4)	4.5 (0.5)	4.8 (1.5)
	Low (24)	4.0 (1.9)	4.3 (1.6)	4.5 (1.8)	4.2 (1.8)

[a]Reversed question in original questionnaire – results switched to treat high values as good.

The data for Perceived Usefulness and Trust was combined into a single average for each student, and tests were performed on these broad categories rather than at the individual question level. If the broad question was found to have significant differences, then individual questions were checked to help identify which portions of the question may be different.

The first t-test performed compared the Perceived Usefulness of the simulation based on the participant's confidence level. The second t-test performed compared the Perceived Usefulness of the simulation based on whether the simulation the participant was included in was correct or not. The results are found in Table 29.

The next t-tests performed compared the Trust shown in the simulation based on the participant's confidence level. The final t-test for this section compared the Trust shown in the

Table 29

t-tests for questionnaire questions related to Automation Complacency

Question	Grouping	Category	n	t-statistic	df	p
Perceived Usefulness	Confidence	High, Low	31, 3	0.95	32	.348
	Correct	Yes, No	17, 17	-0.17	32	.865
Trust	Confidence	High, Low	31, 3	0.26	32	.798
	Correct	Yes, No	17, 17	3.61	29.17	**.001**[a]**

[a] Levene's test for Homogeneity of Variance was violated; since the variances are not the same, these tests were performed using Welch's Two Sample t-test.
** $p<.01$

simulation based on whether the simulation the participant was placed in was correct or not. The results are shown in Table 29.

It is important to note that the only significant difference found was that participants using the correctly working simulation were found to have significantly higher trust than the participants using the incorrectly working simulation.

Further t-tests were performed to determine if there were differences between whether the participant had a correctly working simulation or not based on the separate subquestions under trust. The key words here relate to the words that were part of the question as asked to the participants, including the negative wording for the reversed questions. These subquestions relate to whether the participant felt the simulation was competent in helping them solve the simulation, benevolent in how it tried to help them, low in integrity in helping them, and whether they felt the simulation was trustworthy overall. The results are shown in Table 30. These tests show that participants who were given a correctly working simulation felt that the simulation was more competent, had more integrity, and was more trustworthy.

ANOVAs were performed for all groups with more than two groupings, including treatment groups, self-declared trust levels, and self-declared experience. These ANOVAs were performed to compare both Perceived Usefulness and Trust against the treatment group, the participant's trust level, and the participant's previous experience. The results are shown in Table 31.

For Perceived Usefulness, there was a difference found based on participant's previous experience. Tukey's HSD was run to identify which groups were different. Differences were found between where those with very high previous experience felt the simulation had more perceived usefulness than those who had low previous experience (p=.034).

Further ANOVAs were performed to verify if these differences existed based on if the simulation helped them solve the question more quickly, decreased their solving effectiveness, decreased their solving efficiency, and if the simulation was useful overall for solving truss

Table 30

t-tests for trust and the correctness of the simulation

Subquestion	n	t-statistic	df	*p*
Competent	17, 17	3.062	32	**.004****
Benevolent	17, 17	0.434	32	.667
Low Integrity[a]	17, 17	4.010	24.17	**<.001*****
Trust Overall[a]	17, 17	3.611	29.17	**.001****

[a]Levene's test for Homogeneity of Variance was violated; since the variances are not the same, these tests were performed using Welch's Two Sample t-test

$^{**}p<.01$. $^{***}p<.001$

Table 31

ANOVAs for Perceived Usefulness and Trust

Question	Category	df	F	η	*p*
Perceived Usefulness	Group	3, 30	1.06	.096	.379
Perceived Usefulness	Trust	2, 31	0.94	.057	.403
Perceived Usefulness	Experience	2, 31	4.70	.233	**.017***
Trust	Group	3, 30	5.80	.367	**.003****
Trust	Trust	2, 31	109.57	.876	**<.001*****
Trust	Experience	2, 31	0.62	.039	.542

$^{*}p<.05$. $^{**}p<.01$. $^{***}p<.001$.

problems. In the case of the second and third questions, the data were inverted to show increases rather than decreases, as the questions were deliberately reversed to catch participants who were simply entering answers. The results of these ANOVAs can be found in Table 32.

Despite finding a difference when averaged, no individual question showed a significant difference between experience groups, although there was a potential difference (where *p* was between .05 and .1) between groups where some students potentially found that the simulation decreased their effectiveness.

The ANOVAs also found a significant difference based on Trust and which group the participants belonged to. Tukey's HSD post hoc tests found that the mean of Trust was significantly different between participants in the control group and the deceived group (p=.019), and the informed group and the deceived group (p=.010). No other groups were found to be significant. The deceived group was found to have less trust in the simulation than those in the control and informed groups.

Further ANOVAs were run to verify if any of the subquestions from Trust were also significant. These subquestions include believing the simulation was competent, benevolent, has low integrity, and could be trusted overall. Note that the third question was a reversed question, so the data was inverted to align with the other questions prior to analysis. The results of the ANOVAs are found in Table 33.

Table 32

Specific ANOVAs for Perceived Usefulness Subquestions based on Previous Experience

Subquestion	df	F	η	p
More Quickly	2, 31	0.05	.003	.948
Decrease Effectiveness	2, 31	2.82	.154	.075
Decrease Efficiency	2, 31	1.30	.077	.288
Useful Overall	2, 31	0.12	.008	.887

Table 33

Specific ANOVAs for Trust Subquestions based on Treatment Group

Question	df	F	η	p
Competent	3, 30	3.54	.262	**.026***
Benevolent	3, 30	0.43	.041	.733
Low Integrity[a]	3, 30	5.37	.349	**.004****
Trust Overall	3, 30	4.21	.296	**.013***

[a]Levene's test for Homogeneity of Variance was violated; since the variances are not the same, these tests were performed using Welch's Two Sample t-test.
$^{*}p<.05$. $^{**}p<.01$

Tukey's HSD test found significant differences on the subquestion about the competence of the simulation between the control and misled groups (p=.030). No other comparisons were significant for this subquestion. Tukey's HSD also found significant differences on the subquestion about the simulation having low integrity between the control and misled groups (p=.020). No other comparisons were significant for this subquestion. Tukey's HSD finally found significant differences on the subquestion of overall trust in the simulation between the informed and misled groups (p=.042). All of these differences were between treatment groups with a correctly working simulation and the misled group, which had an incorrectly working simulation while also being informed it might have issues.

The original ANOVAs found one final difference based on the trust rating of the participants and the results of the Trust questions in the Reliance on Technology questionnaire. Tukey's HSD tests showed a significant difference between those with low trust and high trust (p<.001) and between those with low trust and neutral trust (p=.039). The post hoc tests did not find a difference between high and neutral trust.

Further ANOVAs were run to verify which questions of trust might also be different between the different trust groups. These subquestions include believing the simulation was competent, benevolent, and has low integrity. Note that the third question was a reversed question, so the data was inverted to align with the other questions prior to analysis. The overall trust in the simulation was not tested, as this value was used to identify the three groups. This forces the groups to have different means, making the tests redundant. The results of the ANOVAs is found in Table 34.

Tukey's HSD test found significant differences on the competence of the simulation between those with high and low trust (p<.001). Tukey's HSD test also found significant differences on the integrity of the simulation between those with low and high trust (p=.010). No other comparisons were significant.

Qualitative Coding of Questionnaire and Structured Interviews

The structured, open-ended questions and the interviews went through two rounds of coding using the a priori codes shown in Table 9. These codes were approved by the five members of the dissertation committee for use in evaluating the questions, and Table 10 and Table 11 show how each code is assigned to each question. The researcher and another graduate student met to review the definitions of the code prior to coding the questionnaires and interviews. Eight (8) participants were selected for the coders to review, one for each simulation group from each semester. This accounted for 23.5% of the questionnaire responses being coded by both individuals. Due to the low number of interviews, all four interviews were coded by both participants. The interrater agreement for the questionnaires was 94.5% while the interrater

Table 34

Specific ANOVAs for Trust Subquestions based on Trust

Question	df	F	η	p
Competent	2, 31	13.56	.467	**<.001*****
Benevolent	2, 31	0.06	.004	.942
Low Integrity	2, 31	5.13	.249	**.012***

$^{*}p$<.05. $^{***}p$<.001.

agreement for the interviews was 83%. Following meeting and discussing the results, the researcher coded the rest of the questionnaires. The information related to automation complacency, as shown in Table 10, is presented in the remainder of this section while the information related to automation bias, as shown in Table 11 is shown in the Qualitative Coding of Questionnaire and Interviews portion of the next section.

In the process of evaluating the data, it became apparent that some questions addressed a different code than was originally intended with the a priori codes. As such, these codes were added to the relevant question and will be addressed in the reports below. All new codes still came from the existing list. An example is "Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?" that was aimed to show students were being Vigilant in their analysis of the simulation, and a few students instead made comments more in line with simply being Aware of the state of the simulation rather than actively attempting to ensure that it was correct. These additional codes will be indicated in the relevant tables. Table 35 shows each question, the a priori and emergent codes, the number of students who were designated as showing that code, and the percentage of students from that group to do so.

Table 35

Qualitative Coding of Automation Complacency Questions, All Participants

Question	Coding	Count	Percentage
Did you need to recheck your answer because the program said it disagreed with you?	Complacency	30	88.2%
	Vigilance	2	5.9%
Did you change your answer when given the opportunity to do so by the simulation? If so, why?	Switch	29	85.3%
	Complacency[a]	1	2.9%
	Vigilance[a]	4	11.8%
How did you check your answer before submitting it?	Vigilance	1	2.9%
	Complacency[a]	28	82.4%
	Switch	3	8.8%

Question	Coding	Count	Percentage
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Complacency	27	79.4%
What did you do if you disagreed with the simulation?	Vigilance	9	26.5%
	Switch	10	29.4%
	Complacency[a]	12	35.3%
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Vigilance	21	61.8%
How does knowing about the intentional errors impact your thoughts on the simulation?	Complacency	11	32.4%
	Vigilance	10	29.4%
	Switch	1	2.9%
How do you feel about having the simulation try to correct your answers now that you know it may have compared your answers with an incorrect one?	Complacency	23	67.6%
	Vigilance	4	11.8%
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Vigilance	2	5.9%
	Switch	1	2.9%
How would you ensure that a guided simulation or other technology you are using isn't faulty?	Vigilance	19	55.9%
Tell me about yourself, including your use of technology in your learning and with solving trusses using the Method of Joints.	Complacency	4	100%
Tell me about your experience using the guided simulation.	Complacency	1	25%
	Vigilance	0	0%
How often did the simulation disagree with your answer for a force when you first entered it?	Complacency	3	75%
	Vigilance	2	50%
	Switch[a]	1	25%
When the simulation disagreed with your answer, what did you do?	Switch	2	50%
	Complacency	2	50%
	Vigilance	2	50%
How did you feel when you found out the guided simulation might intentionally contain errors?	Complacency	2	50%
	Vigilance[a]	1	25%

[a]Code was an emergent code added while analyzing the data.

Further analysis was performed to see where differences may have existed between the participants based on if they were placed into a correctly working simulation, which simulation group they were in, what their self-reported confidence level is, what their self-reported trust level is, and what their self-reported experience level is. The results of this analysis can be found in Appendix D.

Factors Leading to Automation Bias

This section continues the previous section's attempt to answer the final research question. In particular, this section will specifically address the subquestion of "What factors explain the impact that automation bias has on a participant's ability to recognize errors in that simulation?"

Descriptive Statistics of Reuse Intention and Process Similarity

Means and standard deviations for each subquestion found in the Reuse Intention and Process Similarity questions in the Reliance on Technology questionnaire were calculated and are presented below in the table in this form, mean (SD). As with the previous section in helping analyze the data, the questions related to Reuse Intention and Process Similarity were averaged together for each participant. The results for both averaged results are found in Table 36, while individual question results are found in Table 37 for Reuse Intention and Table 38 for Process Similarity.

Table 36

Descriptive Statistics for Reuse Intention and Process Similarity

Category	Subcategory	Reuse Intention	Process Similarity
Overall	(34)	4.0 (1.8)	4.4 (1.2)
Group	Control (9)	5.2 (1.1)	4.8 (1.0)
	Deceived (9)	3.9 (2.0)	4.4 (1.4)
	Informed (8)	4.3 (1.3)	4.3 (1.4)
	Misled (8)	2.4 (1.6)	4.0 (1.0)
Correct Simulation	Yes (17)	4.7 (1.2)	4.6 (1.2)
	No (17)	3.2 (2.0)	4.3 (1.2)
Confidence	High (31)	4.1 (1.8)	4.5 (1.2)
	Low (4)	2.7 (1.7)	3.9 (1.6)
Trust	High (20)	4.6 (1.3)	4.6 (1.1)
	Neutral (3)	2.6 (1.9)	4.3 (1.3)
	Low (11)	5.2 (0.3)	3.8 (1.1)
Experience	Very High (4)	3.6 (2.2)	4.2 (1.4)
	High (6)	3.9 (1.6)	4.5 (1.1)
	Low (24)	4.0 (1.8)	4.2 (1.6)

Table 37

Descriptive Statistics for Reuse Intention's Subquestions

Category	Subcategory	Task	Same Truss	Similar Truss	Future Truss	Willing Reuse
Overall	(34)	3.4 (1.7)	3.4 (1.7)	4.2 (2.1)	4.5 (2.0)	4.4 (1.9)
Group	Control (9)	4.6 (1.2)	4.6 (1.5)	5.4 (0.9)	5.8 (0.8)	5.4 (1.2)
	Deceived (9)	3.4 (1.8)	3.4 (1.8)	4.3 (2.4)	4.3 (2.1)	4.1 (2.1)
	Informed (8)	3.4 (1.5)	3.5 (1.3)	4.2 (1.8)	5.1 (1.5)	5.1 (1.1)
	Misled (8)	2.0 (1.3)	1.9 (1.4)	2.5 (2.0)	2.6 (2.1)	2.9 (2.0)
Correct Simulation	Yes (17)	4.0 (1.5)	4.1 (1.5)	4.9 (1.5)	5.5 (1.2)	5.3 (1.2)
	No (17)	2.8 (1.7)	2.7 (1.8)	3.5 (2.4)	3.5 (2.2)	3.5 (2.1)
Confidence	High (31)	3.5 (1.7)	3.5 (1.7)	4.3 (2.0)	4.6 (2.0)	4.5 (1.9)
	Low (4)	2.3 (1.5)	2.3 (1.5)	2.7 (2.1)	3.0 (1.7)	3.0 (1.7)
Trust	High (20)	3.8 (1.5)	3.9 (1.6)	4.8 (1.6)	5.2 (1.3)	5.0 (1.2)
	Neutral (3)	4.0 (1.0)	4.0 (1.0)	6.0 (1.0)	6.0 (0.0)	6.0 (1.0)
	Low (11)	2.5 (1.8)	2.3 (1.7)	2.5 (2.0)	2.7 (2.1)	2.8 (2.1)
Experience	Very High (4)	3.0 (2.2)	3.2 (1.9)	3.8 (2.8)	4.0 (2.4)	4.2 (2.2)
	High (6)	3.3 (1.5)	3.3 (1.5)	3.8 (1.8)	4.7 (2.0)	4.3 (1.6)
	Low (24)	3.5 (1.7)	3.4 (1.8)	4.3 (2.1)	4.5 (2.0)	4.5 (2.0)

Table 38

Descriptive Statistics for Process Similarity's Subquestions

Category	Subcategory	Solving Style	Order	Final Answer
Overall	(34)	4.0 (1.6)	4.6 (1.6)	4.6 (1.5)
Group	Control (9)	4.3 (1.3)	5.2 (0.8)	4.8 (1.4)
	Deceived (9)	3.8 (2.0)	4.9 (2.1)	4.7 (1.9)
	Informed (8)	4.0 (1.7)	4.1 (1.8)	4.9 (1.2)
	Misled (8)	3.9 (1.6)	4.1 (1.2)	4.1 (1.6)
Correct Simulation	Yes (17)	4.2 (1.5)	4.7 (1.4)	4.8 (1.3)
	No (17)	3.8 (1.7)	4.5 (1.8)	4.4 (1.7)
Confidence	High (31)	4.0 (1.6)	4.7 (1.6)	4.7 (1.5)
	Low (4)	4.0 (1.7)	4.0 (1.7)	3.7 (1.5)
Trust	High (20)	4.3 (1.4)	4.8 (1.3)	4.6 (1.4)
	Neutral (3)	3.3 (1.5)	2.7 (2.1)	5.3 (0.6)
	Low (11)	3.6 (2.0)	4.9 (1.8)	4.4 (1.9)
Experience	Very High (4)	4.0 (1.8)	4.5 (1.7)	4.0 (1.8)
	High (6)	4.0 (1.7)	3.7 (1.6)	4.8 (1.2)
	Low (24)	4.8 (1.2)	4.7 (1.6)	4.0 (1.8)

t-tests and ANOVAs

In maintaining the same analysis as before, where only two categories exist within a grouping, such as whether the simulation used was correct or not or the self-declared confidence of the participant, then t-tests were used, but where more than two categories exist, such as treatment group, self-declared participant trust or participant experience, then ANOVAs were used. The control and informed groups were combined into a correctly working simulation group, while the deceived and misled groups were combined into an incorrectly working simulation group to allow for broad analysis of the impact of the simulation's responses.

The data for Reuse Intention and Process Similarity were combined for each question separately to provide an overall analysis of the question. If the broad question showed that significant differences existed, then that question was separated and evaluated to determine if a particular question was significantly different.

Reuse Intention was evaluated first, looking to see if different groups were more likely to reuse the simulation. t-tests were performed on Reuse Intention, separating for both the confidence level and the grouping of the simulation based on simulation correctness in the question responses. The results are shown in Table 39.

The next t-tests were performed on Process Similarity, looking to see if there were differences in confidence level and broad grouping of the simulation in the question responses. The results are shown in Table 39.

The only significant difference exists between the correctly working simulation and the incorrectly working simulation. Those participants who were given a working simulation were more likely to reuse the simulation than participants who were not given a working simulation.

Further t-tests were performed to determine which subquestions of Reuse Intention were significantly different between the participants with correctly working simulations and those without. The subquestions researched are whether the participant will reuse the simulation for the same solving task, for the same truss, for similar truss structures, solving future problems, and willing to reuse the simulation again. The results are shown in Table 40.

Those participants with a correctly working simulation showed that in for each subquestion, they were significantly more likely to reuse the simulation than those who were given an incorrectly working simulation.

Following the t-tests, ANOVAs were run to determine if differences existed for each question based on the treatment group, participant trust, and participant experience. Each of these groupings has more than three categories. The results are shown in Table 41.

The treatment group and the participant trust both showed significant differences in relation to reuse intention, while process similarity failed to show any significant differences.

Tukey's HSD tests showed that participants in the control group were significantly more likely to reuse the simulation than participants in the misled group (p=.005), but that no other groups had any differences. Tukey's HSD tests also showed that participants with high trust were also

Table 39

t-tests for Reuse Intention and Process Similarity

Question	Grouping	Category	n	t-statistic	Df	p
Reuse	Confidence	High, Low	31, 3	1.34	32	.190
Intention	Correct[a]	Yes, No	17, 17	2.76	26.78	**.010****
Process	Confidence	High, Low	31, 3	0.59	32	.610
Similarity	Correct	Yes, No	17, 17	0.76	32	.450

[a]Levene's test for Homogeneity of Variance was violated; since the variances are not the same, these tests were performed using Welch's Two Sample t-test.
$^{**}p<.05$.

Table 40

t-tests based on Simulation Correctness and Reuse Intention Subquestions

Subquestion	n	t-statistic	df	p
Same Task	17, 17	2.26	32	**.031***
Same Truss	17, 17	2.43	32	**.021***
Similar Truss[a]	17, 17	2.09	26.50	**.046***
Future Truss[a]	17, 17	3.23	24.60	**.003****
Willing Reuse[a]	17, 17	3.03	24.97	**.006****

[a]Levene's test for Homogeneity of Variance was violated; since the variances are not the same, these tests were performed using Welch's Two Sample t-test.
$^{*}p<.05$. $^{**}p<.01$.

Table 41

ANOVAs for Reuse Intention and Process Similarity

Question	Category	df	F	η	p
Reuse Intention	Group	3, 30	4.72	.321	**.008****
Reuse Intention	Trust	2, 31	7.16	.316	**.003****
Reuse Intention	Experience	2, 31	0.08	.005	.920
Process Similarity	Group	3, 30	0.53	.051	.663
Process Similarity	Trust	2, 31	0.63	.039	.541
Process Similarity	Experience	2, 31	0.29	.018	.752

$^{**}p<.01$.

significantly more likely than those with low trust to reuse the simulation (p=.004). No other differences were found between trust groups.

It should be noted that none of the tests showed any significant differences for Process Similarity. That is, whether participants found the process of solving trusses using the method of joints did not depend on any of the groupings used in this study.

Further ANOVAs were run to determine if there were differences in the subquestions related to Reuse Intention for both treatment group and participant trust. The subquestions researched are whether the participant will reuse the simulation for the same solving task, for the same truss, for similar truss structures, solving future problems, and willing to reuse the simulation again. The results are shown in Table 42.

There were significant differences found between treatment groups for reuse intention of the simulation for: the same task, the same truss, a similar truss, future trusses, and in being willing to use the simulation in the future. Tukey's HSD tests showed that participants in the control group were more significantly likely to reuse the simulation for the same task than the

Table 42

ANOVAs for Reuse Intention Subquestions based on Treatment Group and Participant Trust

Grouping	Subquestion	df	F	η	p
Treatment Group	Same Task	3, 30	4.18	.295	**.014***
	Same Truss	3, 30	4.43	.307	**.011***
	Similar Truss[a]	3, 30	3.59	.264	**.025***
	Future Truss[a]	3, 30	5.48	.354	**.004****
	Willing Reuse[a]	3, 30	3.93	.282	**.018***
Trust	Same Task	2, 31	2.73	.150	**.081**
	Same Truss	2, 31	3.89	.200	**.031***
	Similar Truss	2, 31	7.76	.334	**.002****
	Future Truss[a]	2, 31	10.46	.403	**<.001*****
	Willing Reuse[a]	2, 31	9.07	.369	**<.001*****

[a]Levene's test for Homogeneity of Variance was violated; since the variances are not the same, these tests were performed using Welch's Two Sample t-test.
$^{*}p<.05$. $^{**}p<.01$. $^{***}p<.001$.

misled group (p=.007), participants in the control group were significantly more likely to reuse the simulation to solve the same truss than those in the misled group (p=.006), participants in the control group were significantly more likely to use the simulation to solve a similar truss than those in the misled group (p=.014), that participants in the control group were significantly more likely to reuse the simulation in the future than those in misled group (p=.003), that participants in the informed group were significantly more likely to reuse the simulation in the future than those in the misled group (p=.027), and that participants in the control group were significantly more willing to reuse the simulation than those in the misled group (p=.018).

There were significant differences found between participant trust levels for the reuse intention of the simulation for: the same truss, a similar truss, future trusses, and in being willing to use the simulation in the future. Tukey's HSD test found that participants with high trust were more likely to reuse the simulation to solve the same truss than those with low trust (p=.029), that participants with high trust were more likely to reuse the simulation to solve a similar truss than those with low trust (p=.005), that participants with high trust were more likely to reuse the simulation than those with low trust (p<.001), that participants with high trust were more willing to reuse the simulation than those with low trust (p=.001), and that participants with neutral trust were more willing to reuse the simulation than those with low trust (p=.010).

Qualitative Coding of Questionnaire and Structured Interviews

The coding which took place to answer the factors contributing to automation complacency was also used to answer the factors contributing to automation bias using the a priori codes shown in Table 9. These codes were approved by the five members of the dissertation committee for use in evaluating the questions, and Table 10 and Table 11 show how each code is assigned to each question. The researcher and another graduate student met to review the definitions of the code prior to coding the questionnaires and interviews. The information related to automation bias, as shown in Table 11, is presented in the remainder of this section while the information related to automation complacency, as shown in Table 10, is shown in the Qualitative Coding of Questionnaire and Interviews portion of the previous section. The 94.5% inter-rater reliability for the questionnaire and the 84% inter-rater reliability for the interviews includes this analysis, including the participants reviewed. Table 43 shows each question, the a prior and emergent codes, the number of students who were designated as showing that code, and the percentage of students from that group to do so.

While coding, three questions which were intended to explore automation complacency ended up being coded with a priori codes based on automation bias. The questions "Did you change your answer when given the opportunity by the simulation? If so, why?", "How would you ensure that a guided simulation or other technology you are using isn't faulty?", and "How often did the simulation disagree with your answer?" had participants surprisingly add the codes trust, awareness, and even help in understanding what the participants said. These codes are included here and are analyzed further in Chapter V.

Table 43

Qualitative Coding of Automation Bias Questions, All Participants

Question	Coding	Count	Percentage
How do you feel about using guided simulations to help learn a topic?	Trust	21	61.8%
	Help	2	5.9%
Did you change your answer when given the opportunity by the simulation? If so, why?[b]	Trust	5	14.7%
How did you recheck your answer before resubmitting it?	Trust	29	85.3%
	Help[a]	2	5.9%
How often did you use the help provided by the simulation?	Help	21	61.8%
Did the help provided by the simulation help you solve the truss problem? If so, how?	Help	19	55.9%
	Trust	17	50%

Question	Coding	Count	Percentage
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Awareness	12	35.3%
	Trust[a]	3	8.8%
What did you do if you disagreed with the simulation?	Awareness	0	0%
	Trust	4	11.8%
	Help[a]	2	5.9%
How do you feel about the fact that the purpose of the study was not shared from the beginning?	Trust	1	2.9%
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Awareness	2	5.9%
	Trust[a]	11	32.4%
How does knowing the information above about the intentional errors impact your thoughts on the simulation?	Trust	16	47.1%
Has your perception of the help provided by the simulation changed after learning the true nature of the study and the simulation?	Help	6	17.6%
	Trust	5	14.7%
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Trust	19	55.9%
	Awareness	1	2.9%
How does this study impact your perceived usage of guided simulation moving forward?	Trust	8	23.5%
	Awareness	0	0%
	Help[a]	1	2.9%
When is the ideal time to know that a guided simulation or other technology might be faulty?	Awareness	0	0%
How would you ensure that a guided simulation or other technology you are using isn't faulty?[b]	Help[a]	5	14.7%
Tell me about your experience using the guided simulation.	Trust	2	50%
	Awareness	1	25%
What was your strategy in solving the truss problem?	Trust	1	25%
	Help[a]	1	25%
How often did the simulation disagree with your answer?[b]	Awareness	2	50%
When the program disagreed with your answer, what did you do?	Awareness	1	25%
	Trust[a]	2	50%
What was your experience using the help provided by the simulation?	Help	2	50%

Question	Coding	Count	Percentage
How did you feel when you found out the guided simulation might intentionally contain errors?	Awareness	4	100%
	Trust	2	50%

[a]Code was an emergent code added while analyzing the data. [b]Question received emergent code from other

Further analysis was performed to see where differences may have existed between the participants based on if they were placed into a correctly working simulation, which simulation group they were in, what their self-reported confidence level is, what their self-reported trust level is, and what their self-reported experience level is. The results of this analysis can be found in Appendix E.

CHAPTER V

DISCUSSIONS, CONCLUSIONS, IMPLICATIONS, AND RECOMENDATIONS

This section is divided into three sections. The first section will discuss the findings found in the previous chapter and will provide conclusions. The second will discuss the implications of this study. The third will provide recommendations for future research. The participants were divided across several different metrics to allow for the analysis of automation bias and complacency. First, each participant was placed into a treatment group. Those in the control group received a correctly functioning simulation and weren't informed of potential errors. Those in the deceived group were given a faulty simulation and weren't informed of potential errors. Those in the informed group were given a correctly functioning simulation and were warned that there are potential errors. Those in the misled group were given a faulty simulation and were also warned that there are potential errors. These groups were also combined to correctly working simulations (the control and informed groups) and faulty simulations (the deceived and misled groups). Second, participants were divided on self-reported measures of confidence, trust, and prior experience. This allowed for analysis based on these metrics.

Discussions and Conclusions

Research Question #1: How are the participant's level of automation complacency and the correctness of the simulation that participant is using related?

Automation complacency is measured by determining the extent to which the participants listened to and followed the advice of the simulation. It is about whether the participant performs as instructed by the simulation. Those with a correct simulation should exhibit more changes from Wrong to Right switches, as their answer is checked against a correct force and the advice given by the simulation helps them see the correct way to solve the problem. These participants receive positive feedback from the simulation, influencing their self-regulated learning (Zimmerman & Moylan, 2009). In particular, the feedback provided by the simulation impacts the self-reflection phase the most, influencing future forethought and performance phases. After receiving feedback of a correct answer from the simulation, the participant will have positive self-judgment and self-reaction, which should lead to better task analysis in future cycles. After receiving feedback of an incorrect answer from the simulation, the participant will be able to perform correct self-judgment and self-reaction, beginning to make the necessary changes to get more right answers in the future. This should also lead to better task analysis in the future cycles.

On the other hand, those with a faulty simulation should exhibit more changes from Right to Wrong switches, as their answer is checked against an incorrect force and the advice given by the simulation leads them to an incorrect solution. Their self-regulated learning will be influenced to perform in incorrect ways – their self-judgment will be influenced by the incorrect simulation, leading to a self-regulated learning cycle built upon incorrect information. Some students with faulty simulations were able to self-reflect correctly despite having negative feedback and were able to find the correct answer regardless of what the simulation gave them, indicating a positive self-regulated learning pattern regardless of what the simulation demonstrated.

For all participants, regardless of treatment group, 63.9% of the switches were Right to Right (RR) switches, meaning the original answer was correct and did not change (for all rates referenced, see Table 15). Those in the control group had a 69.8% and in the informed group had a 75% RR switch rate, while those in the deceived group had a 50% RR switch rate and those in the misled group had a 61.6% RR switch rate.

On the other hand, those in the control group had a 0.8% Right to Wrong (RW) switch rate, those in the informed group had a 0.9% RW switch rate, those in the deceived group had a 18.3% RW switch rate, and those in the misled group had a 9.8% RW switch rate (for all rates referenced, see Table 15).

These numbers demonstrate that participants with a correctly working simulation are more likely to find and remain on a correct answer than those who did not have a correctly working simulation. This lines up with findings in the t-tests and ANOVAs, where the differences in the number of RW switches was found to be significant between those with a correctly working simulation and those given the faulty simulation. This is reinforced with those in the deceived group who made significantly more RW switches than those in the control and informed groups. This is what we would expect, where those with a correctly working simulation will be told that a correct answer is actually correct, hopefully suggesting that they not change the answer, while those with a faulty simulation will be told that a correct answer is incorrect, which should cause them to change their answer to accommodate what they are told. Such action is also supported by the self-regulated learning cycle reinforcing the participant's learning. As participants self-reflect on the feedback given by the simulation and plan to participate in future tasks, the feedback received when performing should influence future work throughout the simulation.

Similarly, t-tests showed that those with a faulty simulation were more likely to switch their answer in any direction than those who had a correct simulation. This makes sense, as any wrong answer will be caught by both simulations, but only the faulty simulation will try to get participants to change a correct answer into an incorrect answer. It was expected that most of the changes should happen within the misled and deceived groups, assuming that participants are getting at least some of the problems correct. With many of the misled and deceived groups still able to get correct answers, this suggests that there was some dissonance in the self-regulation of the participants. That is, some participants allowed the simulation to influence their forethought as they planned to solve the forces on additional trusses, changed the way they performed and what help they sought while solving the forces, and then adjusted the reflection they had after getting feedback again. It is possible that some participants adjusted their mental models of the method of joints in either direction based on the feedback they were receiving.

The participant's confidence level did not have a significant difference on whether a user changed their answers, either as RW switches, WR switches, or even WW switches. This is surprising, as it was expected that those with high confidence are more confident in their answers and would make less switches because of that. But, the data did not support this. This may be due to confidence levels shifting over the course of the simulations, particularly if participants started to doubt themselves based on the feedback from a faulty simulation, however, more study would be needed to investigate this hypothesis.

Similarly, there were no differences in switching based on the self-reported trust. As before, one would expect more switches made by those who trusted the simulation to a greater extent, particularly if they had a faulty simulation. This is surprising, since there were more participants with high trust whom had a correct simulation (control and informed groups) and more participants with low trust which had a faulty simulation (deceived and misled groups, see Table 44). The analysis is unsurprising in this area, with those with a faulty simulation ending with a lower trust rating, but this could also show that participants are starting to reject the advice from the simulation, reducing the differences between the different groups.

Table 44

Counts of Participants based on Simulation Group and Trust

		Trust		
		High	Neutral	Low
Group	Control	8	0	1
	Deceived	2	2	5
	Informed	7	1	0
	Misled	3	0	5

There were also no significant differences found based on the self-reported experience of the participants. Given the answers provided to previously used technology, this isn't to surprising, as Table 13 shows that most students considered using their LMS or online homework systems as the simulation tool they had experience with. The differences between LMS's, online homework systems, YouTube videos, and computational engines and a guided simulation are quite stark – despite feeling that they may have high or even very high experience using technology, no participants really had experience with anything like the guided simulation provided here.

These findings support the Theory of Technology Dominance, where those new to a particular problem, or technological solution to that problem, are more likely to rely heavily on the technology than those who aren't. Considering that a greater number of switches occurred more frequently, and in significant numbers, for those in the deceived group and, to a lesser extent, in the misled group, as compared to the control and informed group, indicates that student switching rates are dependent on whether the simulation they have is working correctly or not. These participants behaved exactly as we would have expected participants to, based on the simulation that they were provided with. With data indicating significant differences in both Right to Wrong choices, for those with a faulty simulation making these switches more often, and Wrong to Right choices, for those with a correct simulation making these switches more often, the researcher has concluded that both groups of participants are showing automation complacency. This indicates that participants are performing actions based on the state of the simulation, whether it is correct or not. Further, their self-regulation in their learning adjusts their mental models according to the correctness of their simulation.

Research Question #2: How is automation bias related to a participant's ability to recognize errors in a simulation?

Automation bias is measured by determining the extent to which participants trusted the simulation over their own intuition. Automation bias is shown when a person gives preference to the technology over any evidence to the contrary. In other words, automation bias is shown when the participant trusts the simulation. Whether automation bias can be shown depends partially on the ways in which the participant switches their answers, but it also depends on how their initial and final entered force values match either the correct answers or the answers that their simulation say are correct regardless of whether the solution path used by the simulation is correct or incorrect. For this reason, information related to automation complacency and decision switching impacts the automation bias, but the accuracy of each participant's answers matters, too. The participant's initial accuracy is calculated as the number of correctly entered force values divided by the total number of forces, when we consider only the first answer provided by the participant to the simulation. The participant's final accuracy is calculated as the number of correctly entered force values divided by the total number of forces, counting only the second answer provided by the participant to the simulation.

The control group had a fairly high Right to Right (RR) switch rate of 69.8% and a Wrong to Right (WR) switch rate of 9.5%. The control group had a simulation coaching them towards a correct answer. They also had a minimal Right to Wrong (RW) switch rate of 0.8%.

Their initial accuracy was 49.2%, but rose to 57.9% for final accuracy. This suggests that students in the control group used the simulation appropriately and, since they were given correct feedback and advice, it also suggests that they would trust the simulation and make appropriate changes. Very few members of the control group changed their correct answer to an incorrect answer.

The informed group had an even higher RR switch rate of 75.0%, with a WR switch rate of 13.4% and a RW switch rate of 0.9%. The informed group also had a simulation coaching them towards a correct answer. Their initial accuracy was 68.8%, with a final accuracy of 81.2%. The participants in this group started with higher accuracy and maintained their results throughout the simulation. It appears that this group were more aware of the state of the simulation, and the dangers that might come from mistakes in it, and so were more careful in using the method of joints, thus having higher accuracies than the control group.

The misled group had a decent RR switch rate of 61.6% and a WR switch rate of 11.6% and a RW switch rate of 9.8%. The reader is reminded that the misled group had a simulation coaching them towards an incorrect answer. Their initial accuracy was 58.0%, and their final accuracy was 59.8%. It appears that knowing there might be potential problems helped some of the participants, who were able to find a right answer even after being informed their answer was wrong, but some of the participants still followed the advice of the incorrect simulation and changed correct answers into incorrect answers. This caused their initial and final accuracies to remain roughly the same. It also shows that students are showing some bias towards the simulation rather than themselves.

The deceived group had the lowest RR switch rate of 50% and a decent WR switch rate of 10.3%, but the group had a fairly high RW switch rate of 18.3%. The deceived group had a simulation coaching them towards an incorrect answer. This is confirmed with an initial accuracy of 50% and a final accuracy of 42.1%. This group was the only group to have a drop from initial to final accuracy. This shows the amount of bias exhibited by the deceived group. The main reason to change a right answer to a wrong answer is that, after entering the right answer, the simulation informed them they were wrong and asked them to change their answer.

This supports the Theory of Technology Dominance in that most participants were new to the method of joints having just recently learned and applied it as an analysis technique, and, unless given a strong reason otherwise, they reacted to the feedback and changed their answers accordingly. This suggests that their self-regulated learning takes the feedback in while self-reflecting on the results of a particular joint, impacting their forethought phase as they prepare for a new joint, and finally impacting their performance phase as they find the forces in the new joint. Even the misled group had a relatively high percentage of RW switches, suggesting that despite knowing there might be problems, these participants still made changes based on feedback from the simulation during their forethought and performance phases.

Additionally, the reader is reminded that there was a difference in the simulations between Fall and Spring semester. However, significance tests showed that there were no significant differences in any accuracy between the Spring's simulation, where participants solved an entire truss from beginning to end, and the Fall's simulation, where participants worked on separate joint problems form different trusses. This shows that biases are more inherent in the treatment groups than they are in the simulation itself.

Significant differences were found between those with high confidence and low confidence in both initial and final accuracy. Those who felt more confident in their answers were shown to have higher accuracies than those who weren't confident in their answers. The more confident the participant was, the more likely they were to have a higher initial and final accuracy. This makes sense and aligns with Goddard et al., (2014), as the surer they are, the more likely they are to get future problems right. This also suggests that they are more inclined to be critical of the feedback from the simulation and trust their prior mental models as opposed to the simulation. If they are using the helps during the performance phase and those helps are guiding them to the correct answer, they will self-reflect on those and continue to do so in future forethought and performance phases. If they attempt to use the helps and those helps are guiding them to an incorrect answer, they may begin to ignore the helps and use their own judgment in future forethought and performance phases. Those with high confidence may find more correct answers due to repeated performance, while those with low confidence are continually being asked to recalculate their answer and may struggle with or without the helps during the performance phase.

ANOVAs found that participants in the deceived group had statistically significant lower final accuracy than those in the informed group. Not knowing that the system might have errors impacts the final accuracy of the participant, as those in the deceived group have no reason to stick with a correct answer when the simulation informs them that their answer is wrong even though it is correct. These participants may be struggling in the self-reflection phase, as the feedback is not aligning with their initial answers, especially at first. If they consistently try to use the right method during the performance phase, the feedback from the simulation will continue to negatively impact their self-judgment during the self-reflection phase.

Neither the self-declared trust given to the simulation, nor the previous simulation experience of the participant showed any significant differences for accuracy. This is interesting, because it was expected that there would be a difference based on trust. Those who trust the system should have accuracies more reflective of their treatment group, with those with correctly working simulations receiving higher accuracies than those without correctly working simulations. Likewise, those with more experience should have been able to spot the differences sooner and made adjustments accordingly. One reason this may have happened was that the group sizes for the different experience groups were both not entirely uniform while also being low for the very high experiences, with only four participants. There may not have been enough of a sample size to find significant differences with these sample sizes.

These results show what the research team was expecting. Those with correctly working simulations tended to be guided towards a higher WR switch rate and higher final accuracy, while those provided with warnings that the system might need contain errors were more aware of the state of the simulation and likewise had higher WR switch rates, higher final accuracies, and lower RW switch rates. Those with an incorrect simulation also had higher RW switch rates. The informed group had the best switch rates and better accuracies, suggesting that their awareness of the simulation led to higher performance. All of this aligns with the Theory of Technology Dominance, which suggests that these participants will place higher value on the feedback from the system, as shown in the higher accuracy rates of those with a correctly working simulation and lower rates for those who were not warned their incorrectly working simulation might contain errors (Arnold & Sutton, 1998).

Research Question #3: What factors explain the automation bias and automation complacency that the participants are experiencing?

The answers to the previous questions showed that participants in the study showed automation bias and automation complacency according to the treatment group they were in. In particular, the control and the informed group both had fairly high Wrong to Right (WR) switch rates while both the deceived and misled had higher Right to Wrong (RW) switch rates, showing that, even if they were somewhat aware of the state of the system, the participants showed automation complacency enough to follow the instructions of the simulation and to change their answers accordingly.

The control and informed groups also showed an increase from initial accuracy to final accuracy which fails to deny automation bias – that is, there is no evidence they rejected the advice given by their correctly working simulation. In fact, self-regulated learning suggests that the automation bias is likely to be more at play, as they are constantly receiving positive feedback during their self-reflection phase, influencing their future forethought and performance phases. The misled group, however, maintained a similar level of initial and final accuracy and had a similar number of WR and RW switches within the group – that is, some participants are still showing that they are following the advice of the simulation and changing right answers to wrong answers despite having been told there may be issues with the simulation. Because they were warned of potential errors in the simulation, it is possible that they exercised greater care when performing their self-evaluations and chose to trust themselves more than the simulation. The deceived group showed automation bias in that they had nearly 20% RW switches and their final accuracy dropped from their initial accuracy, showing an automation bias towards the incorrect simulation. They also demonstrate that they are allowing their self-reflection phase to favor the feedback from the incorrect simulation, impacting their future forethought and performance phases.

What is missing from this analysis is the factors and reasons that may lead to this automation complacency and automation bias. This subsection will now attempt to answer each of the two subquestions in order.

Research Question #3a: What factors explain the correlation between a participant's level of automation complacency and the correctness of the simulation that participant is using?

Both the quantitative and qualitative results from the Reliance on Technology Questionnaire were used to identify factors leading to automation complacency. First, the quantitative responses to the Likert scale questions related to the Perceived Usefulness and Trust found in the Reliance on Technology Questionnaire were evaluated to look for factors. Then, the qualitative coding of the free response questions found both before and after the debriefing during the Reliance on Technology Questionnaire were coded to look for factors relating to automation complacency. Finally, the interviews of willing participants following the questionnaire were coded to look for these factors.

Perceived Usefulness. In reviewing Table 26, Table 27, and Table 28, it is clear that those with a correctly working simulation rated the questions pertaining to Perceived Usefulness and Trust higher than those with an incorrectly working simulation. In fact, looking at the treatment groups, for Perceived Usefulness, the control group rated each element higher than the informed group, which was higher than the deceived group, which was higher than the misled group. Within each correctness group, those who were warned about possible errors showed lower perceived usefulness of the simulation. It appears that participants with a correctly working simulation could see the simulation being more useful than those without but that those who were warned of potential errors would find the simulation less useful than their peers with the same type of simulation, whether correct or faulty.

Trust. For trust, things were more complicated. While all of those with a correctly working simulation had a higher trust rating than those without, such a relationship breaks down when you look at the individual components. Participant ratings of the competence of the simulation were similar to perceived usefulness with descending levels for control, informed, deceived, and misled groups. This again suggests that knowing of possible errors impacted how competent the participants viewed the simulation. In terms of the whether the simulation given to the participant was working correctly or not, however, the deceived and misled groups had the same mean, suggesting that knowing about possible errors had no impact on how much the participants trusted the simulation if it wasn't working correctly. One surprise was that those in the informed group had higher overall trust in the simulation than the control group. It is possible that this is due to these participants being warned of errors but, ultimately, not having any. Or perhaps, having been notified that the simulation had an error meant they attributed a more thorough simulation design than what they may normally prescribe to a general simulation they had encountered.

Significant tests provided further evidence. The only test that ended up showing significance for Perceived Usefulness was between those with very high prior experience and those with low prior experience on computer-based tools. In all other tests, there were no differences found. It is probable that because of their previous experience using computer-based tools, those with very high experience came in with better expectations and were able to begin working with the simulation more quickly. It is possible that the experience provided by using LMSs and online homework systems helped to inform these participants with a better preconceived notion about the simulation. Those with very high prior experience also had the fewest number of total decision switches at eight (8) switches, all of which were WR switches, meaning that the only answers they changed went from wrong answers to right answers. Since these participants are getting correct answers, it is likely that they have a higher perceived usefulness of the simulation.

Significant tests found more differences for the Trust questions. Those who had a correctly working simulation rated their trust higher than those who did not, which shows that, at some level, those who were constantly told to rework their answers were more likely to begin to suspect something was wrong and thus may have their trust decrease over time. When looking at the particular subquestions of trust, those with a correctly working simulation stated that they found the simulation to be more competent in solving the truss, had higher integrity (as interpreted by the participant), and had higher trust in it as a tool for solving trusses than those with an incorrectly working simulation. Again, this makes sense, as they have a working simulation that lines up with their understanding of the method. For those with an incorrect simulation, the frustration of constantly being told they are doing it wrong lines up with a decrease in trust and a belief in the incompetence or integrity of the simulation. This is probably even more heightened when an individual feels they are operationalizing the method of joints correctly and are very comfortable with their application of the concept of equilibrium in developing the equilibrium equations. This also falls in line with self-regulated learning theory, as the students are internalizing correct mental models in the self-reflection phase, impacting their forethought phase as they prepare for the new joint and, after the performance phase, they receive positive feedback about their answer, continuing the cycle.

Significant differences were also found for Trust based on the treatment groups. Both the control and the informed groups showed higher overall trust in the simulation than the deceived group. The deceived group is the one that is constantly being forced to rework their solutions while under the impression that the system is working. It was expected that these participants would have lower trust. In terms of specific subquestions for trust, the control group found the simulation to be more competent and to have more integrity than the misled group did. This reinforces the fact that knowing about possible errors coupled with having an incorrect simulation leads to lower overall trust. What is interesting is that only the misled group was significantly different than the control group. The deceived group was not significantly lower for any of these subquestions, indicating that there is something about knowing about potential errors ahead of time that may be factoring in. The fact the informed group also doesn't show up as significantly different further suggests that it is the combination of knowing about potential errors and having an incorrectly working simulation has a higher impact on trusting the simulation than either one alone does. The final significant difference found here was a difference on overall trust between the informed group and the misled group. This suggests that knowing of potential errors ahead of time may cause more vigilance from the participant in the solution process which then reinforces the participant's trust level based on whether the system does or does not work, with working systems creating more trust and failing systems creating less trust.

The final significant differences found for the Trust question involved the self-declared trust groups for the participants. For the overall questions, the participants with low trust were significantly different than those with high and neutral trust. This is not surprising, but it does suggest that those with low trust were farther from neutral trust (a rating of 4) than those with high trust were. Those with low trust specifically found the simulation to be less competent and to have lower integrity than those with high trust. This is also not surprising, especially as the Trust question from the questionnaire was used to create the groupings around trust. What this does strongly show is that the self-declared trust has an impact on most of the subquestions. Interestingly, the question of the benevolence of the simulation was not found to be significantly different between any groups, suggesting that this subquestion may not be relevant for these types of studies.

Free Response Qualitative Analysis. The coding of the free responses to the questionnaire questions focused on the written responses to questions about complacency, vigilance, and switching of answers. These codes were selected a priori based on the works of Billings et al., (1976) and Wandtner (2018). A response that showed that the participant went along with the program was coded as Complacency. A response that showed that the participant was actively questioning the simulation for correctness was coded as Vigilant. A response that showed that the participant actively switched their response was marked as Switch. After comparing the results of the codes overall and based on the participant's simulation correctness, treatment group, self-reported confidence, trust, and experience (see Table 35, Table 45, Table 46, Table 47, Table 48, and Table 49) as well as demographic data (see Table 50, Table 51, and Table 52) the following patterns were found.

Complacency. Prior to the debriefing, most participants (30) said that they checked their answer again because the program asked them to. This shows that, especially at first, participants were mostly complacent. This didn't depend on the correctness of the simulation, as the same percentage (88%) of participants showed this complacency, regardless of the correctness of the simulation. However, every participant in the control and misled groups specifically mentioned rechecking their answers after being asked to by the simulation. The control group has no reason to doubt the simulation, so this result is not surprising. The misled group may be doing so because, if they are correct, the system will actually inform them their answer is wrong. They are likely to listen to the advice and double check their answer. The deceived group (7) and the informed group (6) had a lower percentage of participants, but most still followed the advice of the program and rechecked their answer. This complacency appeared across gender (males (22) and females (8)), age (18 to 21 (15), 22 to 24 (12) and 25 and older (3)), and major (mechanical (17), civil (9), and others (3)).

Twenty-seven (27) participants agreed that the simulation told them their answer was wrong and that they originally agreed with the simulation. This was evenly split amongst the treatment groups, with seven in each of the control, deceived, and informed groups but only six in the misled group. This suggests that regardless of situation, the initial response of participants was to listen to the simulation. There were a few specific comments from students about starting to doubt, but most of this was captured by vigilance instead of complacency. Every student who stated they had low confidence in their own abilities showed that they agreed with the simulation on their answers being wrong, which makes sense if their confidence rating is accurate. The researcher was surprised to find that 9 of the 11 participants with low trust showed that they were complacent at this level of the simulation. In addition, 20 of 25 males and 8 of 9 females; 16 of 18 of those 18 to 21, 10 of 13 of those 21 to 24, and all three of those 25 and older; and 17 of 21 mechanical and all other majors showed that they were complacent at this level. It's important to note that this question came after the self-reported trust question, which was part of the Likert scale questions.

There were a few surprises of participants coded for complacency even though the original question did not ask for that. One student from the informed group specifically commented that they changed their answer specifically because they were told to. In addition, twenty-eight (28) participants specified that they rechecked their answers specifically because they were asked. This happened more for those using a correct simulation (15) than an incorrect simulation (13), with the control group (9) and deceived group (7) having more of their participants showing complacency in checking their answer than the informed (6) or misled (5) groups. These numbers include 20 of 25 males and 8 of 9 females; 13 of 25 of those 18 to 21, 12 of 13 of those 22 to 24, and all of those 25 and older; and 17 of 21 mechanical, 7 of 9 civil, and all other majors. It also appears that those who were informed of potential errors were less likely to check their answer before resubmitting.

Another surprise was found when students were asked what they did if they disagreed with the simulation. Twelve participants specifically mentioned doing what the simulation asked them to do – rechecking their answer and switching it, if necessary. This was evenly split between the correctness of the simulation. Both the informed and control groups had three (3) participants make such a statement, while six (6) of the deceived participants stated that they had done so. None of the misled participants made any statements in this regard. This shows that those who had a correct simulation were complacent about following the advice of the simulations, while those with an incorrect simulation split specifically over whether they were previously informed of potential errors or not. Being told that your work is wrong after being told the system might be wrong seems to cause participants to be less complacent in using the system.

After being informed of the true nature of the study and which group they belonged to, the tune of complacency changed. This is to be expected – participants expressed at this point that they thought something was going on. This only strengthens the Theory of Technology Dominance because these participants still showed that they were complacent! Even if they felt something might have been off, many of them still switched answers and commented on doing so due to the program asking them to. It also shows that some students, particularly those with a faulty simulation, struggled with the self-reflection phase. They were constantly informed that a correct answer was wrong and, if they were confident in their answers and believed they had the right method, they would constantly have to wrestle with the feedback from the simulation and their self-evaluation of their work. This could go one of two ways – either the participant switched their model to either get the "right" answer or to stop the system from telling them they were wrong, both of which affect their forethought phase, or they decided they were right and ignored the feedback and exercising their self-control by simply leaving their answers alone, which would affect their performance and self-reflection phases.

A third of participants (11) commented that they still trust simulations in general. Comments were made that showed that some of these participants were still second guessing themselves as they filled out the questionnaire. There were also notes that the participant's confidence in their abilities came down due to constantly being told their answers were incorrect. Those with a faulty simulation were much more likely to mention their complacency in regards to the simulation than those with a correctly working simulation, seven (7) to four (4). In particular, the deceived group had the highest number of participants show they were complacent in their answer (4), while the misled group was close behind (3) and both the control and informed groups had a small number (2). All three participants who rated themselves as having neutral trust of the simulation specified their complacency in their answers while only about a third of those with high trust (6) or low trust (2) showed complacency in their answers. None of those who stated they had very high experience showed complacency in their answers. The participants who were actively deceived by the simulation itself tended to doubt themselves more than they doubted the simulation and had a bigger revelation about the impact the simulation had on them than did those with a correctly working simulation.

Unlike the previous question, nearly two-thirds of participants (23) showed complacency in their answers when asked how they felt about knowing the simulation might have actively compared their answer to a wrong one. This was fairly evenly split along the correctness of the simulation, with twelve (12) participants provided with a correctly working simulation and eleven (11) participants provided with a faulty simulation showing complacency. The control group had the most participants show complacency (8), followed by the deceived group (6), the misled group (5), and the informed group (4). Both groups that were simply told to use the simulation showed more complacency in their answer than those who were informed there might be errors, which is what was expected. Participants forewarned about potential problems should be less likely to be complacent about using the simulation than those who weren't. One participant stated that, had they known that the simulation might have been incorrect, they would have trusted themselves more. This particular participant was not only using a correct simulation but they were also in the informed group. The demographic breakdown of those that answered this question in a way that included 20 of 25 males and 3 of 9 females; 10 of 18 of those 18 to 21, 11 of 13 of those 21 to 24, and 2 of 3 of those 25 and older; and 14 of 21 mechanical, 6 of 9 civil, and 3 of 4 other majors.

From a demographic perspective, it appears that both males and females showed similar levels of complacency with using the program. The only exception to this was the question about how they responded to possibly being corrected by an incorrectly programmed simulation, where 20 of the 25 males expressed some form of complacency but only 3 of 9 females expressed the same complacency. This suggests that females became less complacent following the debriefing than did their male counterparts.

Age also plays a small factor, with almost all students showing complacency when asked to recheck their answers. However, those 18 to 21 (13 of 18) showed less complacency when asked how they checked their answer than those 22 to 24 (12 of 13) or 25 and older (3 of 3). Those 18 to 21 (16 of 18) showed more complacency when stating that they agreed with the simulation over their own answer compared to those 22 to 24 (9 of 13) and those 25 and older (2 of 3). This suggests that while they didn't want to check their answers just because the program suggested that they should, the lower students were more likely to agree that their answers were wrong. Another reason these might be different is that the older students may have been more willing to check their answer but were less likely to assume that the program was correct.

The interviews further supported these findings. All four interviewees stated that they enjoy using technology to learn. One of the participants, who was in the misled group, stated that they often found themselves simply going with the feedback provided by the system. Three participants reported just following along with the simulation, doing as it asked regardless of their confidence in their answers. Two of the participants specifically mentioned that they would still do what the simulation said even after learning about the true nature of the study. One of these was in the control group, so their experience was with the simulation working correctly. The other was in the deceived group, which came a little as a surprise since they were actively misled by the simulation. Despite this, their previous experience with simulations and their desire to use technology shows that overall, they still want to follow the directions of the simulations they learn from.

Vigilance. One of the hopes of using the simulation was that participants would learn to be less complacent about their use of technology. Unfortunately, vigilance was the code that appeared least often, especially prior to the debriefing in the reliance on technology questionnaire. Prior to the debriefing, when asked if the participants needed to recheck their answer by the simulation, only two students made statements related to vigilance. Both of these participants had an incorrect simulation, with one each from the misled and the deceived group. The fact that seven other participants were informed of potential errors and yet did not indicate vigilance shows the hold that technology has on participants, further supporting the Theory of Technology Dominance – those who are new to a problem type are more likely to trust the technology. It makes sense that experience in solving a particular problem with a known process would help a simulation user identify when a simulation is wrong as it continually contradicts their solutions.

The next question was about whether participants changed their answer when given the opportunity. Four participants gave responses that coded to vigilance, which was a surprise since the question was not intended to look at vigilance. These four showed that they were starting to catch on as they went – they may have been complacent at the beginning, but as time went on, they started to suspect something was going on. Two participants came from the deceived group, while two came from the misled group. All four participants had a faulty simulation, showed low trust in the simulation, had high confidence, and had low prior experience with simulations. This also suggests that they were self-reflective of the problem and of the simulation itself, choosing to use their past experience rather than the guidance of the simulation as they analyzed the task during the forethought phase.

When asked how participants checked their answer before resubmitting to the simulation, only one participant from the misled group mentioned actively thinking the simulation might be wrong. This participant has high confidence, low trust, and low experience, but they were forewarned about potential errors in the simulation. It is telling that this question was specifically looking for responses related to vigilance and yet most (28) participants responded with answers that were more in line with complacency, as described above.

Nine participants did express concern about constantly being told their answers were wrong, showing some form of vigilance. Only one from each of the control and informed groups indicated an increase in vigilance, including the participant who felt that the simulation was not working correctly described under complacency above, while five of the misled participants and two of the deceived participants indicated a vigilant attitude. The participants with a faulty simulation and knowledge of potential problems should have been the highest group, while those who had a faulty simulation without that knowledge only had a few participants figure out that they needed to watch the simulation more closely.

The next set of questions came after the debriefing. The most interesting thing to note is the sudden rise in answers that showed vigilance. When asked whether the participant suspected that anything could be wrong with the simulation, twenty-one (21) participants said that they suspected something was wrong. What is surprising is that eight (8) of the participants with a correctly working simulation thought something was up, including three from the control group. The five from the informed group is not as surprising, as they were warned, but the fact that they actually had a correctly functioning simulation should have helped them better rely on the simulation. This indicated that these participants may have trusted their own work too much, as shown by the participant whose answers are consistent with those with a faulty simulation despite having a correct one. The thirteen with a faulty simulation were constantly told that correct answers were actually incorrect. The surprise here is that seven of the deceived group said they were suspicious of the program. This is one more than the number in the misled group which should have known something was wrong fairly quickly because they had knowledge that

that errors might be present. Still, participants voiced most of their suspicions only after the debriefing occurred. It could be that they are only trying to claim that they were suspicious after learning the truth and weren't prior to that.. Regardless of the specific level of suspicion and where it started, it is clear that most students still gave priority to the program. Most students did change their answers, with many Right to Wrong changes made amongst the deceived and misled groups. So, even if they were suspicious, these participants still followed the simulation's advice. This directly aligns with the Theory of Technology Dominance and is expected of those with low levels of experience with the subject matter and the decision tool (Arnold & Sutton, 1998). A similar finding was also shown to affect new graduate student users of the Web of Science, a data base of scholarly articles that assisted in finding related articles, who sometimes felt that the system did not find all of the relevant articles but did not look further than the provided articles from the Web of Science (Lou & Sun, 2021).

The data also showed that, despite the experience of working through the simulation, only ten (10) participants indicated that they would be more vigilant in the future. A few comments intimated that they felt the study was clever, but they'd only have to be vigilant when in a similar type of study, not when engaging in learning from simulations in general. The informed group had the most participants who mentioned specifically looking at other simulations and being more vigilant with them (4), followed by the deceived group (3), the misled group (2), and the control group (1). The misled group and the deceived group should have been more willing to apply their vigilance outside of the study itself, as they had the most active harm being done to them. It is interesting that the informed group had the most participants actively applying this beyond the study. The fact that they were warned of potential errors but then received correct advice and checking may have helped them be more aware of the problem in general, while those who received a faulty simulation of any type may have been too fixated on their specific personal experience rather than looking to expand that experience to other simulations in the future. It is possible, that those with a faulty simulation may approach a future simulation with more vigilance than they indicated in their questionnaire answers.

This is followed by a low number of vigilance responses when asked about knowing that the simulation was comparing their answers to incorrect answers. Two each came from the misled and deceived groups, those most impacted by the negative responses, and all four were in the high confidence group and had either neutral or low trust (evenly split between the trust levels). Again, it is surprising that these numbers are not higher – it appears that, like the previous question, most participants seem to think this type of vigilance is only necessary when the simulation is being used to measure the participants' reliance on technology.

Further, only two participants commented on the process of using the simulation and its use of potential errors, and how that impacted their vigilance moving forward. Most tied their answers right in to this particular study, with only one from the control group and one from the misled group indicating any changes in how they would use simulations moving forward.

When the question was changed to how they would ensure that a guided simulation was working correctly in the future, the number expressing vigilance grew to nineteen (19). This was almost evenly split between each of the treatment groups, with five in each group except the deceived group which had only four. This indicates that most participants are still thinking about their interaction in the study in terms of the study until they are specifically asked to think about the potential future impact of their experience.

As for the interviewees, all four specifically mentioned vigilance during the initial discussion of their working in the simulation. However, only those in the misled group indicated that they were vigilant when asked deeper questions, and only one of those two specifically mentioned that they should continue to be vigilant when asked about how their reaction when they found out the true purpose of the study.

Switch. The greatest indication of complacency is whether participants switched their answers based on feedback from the simulation. This was already shown to be taking place during the analysis of simulation results, where the number of right to wrong changes was high amongst the deceived and misled groups and Wrong to Right switches were high amongst the control and informed groups. This is backed up by participant responses on the qualitative portion of the Reliance on Technology questionnaire.

Twenty-nine participants specified that they changed their answers when prompted by the simulation. This was evenly divided between the correct and faulty simulations. Every member of the deceived group stated that they changed their answer, while all but one of the control group did the same. Only six in each of the informed and misled groups did the same. While a majority of the participants changed their answers, it should be noted that a higher percentage of the participants that were not warned that there could be errors stated they changed their answers. This is likely due to an assumption these participants made during the forethought phase about the simulation being built to help them learn and practice the Method of Joints. They may have assumed, especially early on, that they were making mistakes and that the simulation was catching them, leading them to make more changes than those who were given warning of potential errors in the simulation, who did not necessarily make the same assumption.

All participants who stated they had low confidence changed their answers when given the chance, which makes sense, but so did twenty-six of those with high confidence. Even if participants are confident in their responses, there still seems to be a tendency to rely on the simulation when making their choices, especially as they started working with it. The level of self-reported trust also showed a large proportion of students who made changes regardless of their individual trust level. All four participants with very high experience with simulations stated that they changed their answers when asked, while four of six of those with high experience and twenty-one of twenty-four of those with low experience with simulations stated they changed their answer. There were higher percentages for those with previous experience than those with low experience, but the rates were still high. Again, this parallels the findings found in switch rates as shown in the discussion above.

Three participants added responses that were coded as switching when asked how they rechecked their answers. This wasn't looked for, but those who did so specifically mentioned that they simply switched their answers due to getting told that their answers were right but of the wrong sign. In other words, they had the right work, but the simulation disagreed on whether the force was in tension or compression. These changes only occurred with one person each from the deceived, informed, and misled groups. Interestingly, one of these was following correct advice and was working on the problem from the wrong side while the other two were working on the simulation correctly and chose to switch their process of solving the truss as informed by the simulation. This also shows that these participants were engaged in self-judgment during the self-reflection phase of learning.

When asked what the participants did when the simulation disagreed with their answer, ten participants mentioned that they specifically changed their answer because they were asked to do so by the simulation. Seven of these participants had an incorrect working simulation while three had the correctly working simulation. Four of these participants were in the deceived group, while three were in the misled group, two in the informed group, and only one in the control group. What's striking here is that the control group had fewer people who specifically stated that they switched their answer than those in the informed group. The control group were not told there might be potential errors, and yet only one of their number mentioned switching their answer even though they disagreed with the simulation. It was expected that those in the informed group, who were told there might be errors, would be less likely to change their answers when they disagreed with the simulation. This effect is likely to be specific to the participants in this study, but it would be worth investigating further. Another surprise is that eight of the ten participants expressing that they switched their answer despite disagreeing with the simulation had high confidence, while only two of those with low confidence expressed the same reliance on the simulation. It was expected that the low confidence participants would be more likely to switch just because, since their lower confidence implies that they are more reliant on the simulation to help them get through the system.

After being debriefed, there were few expectations that participants would comment on switching their answer. At that point, the researcher was more interested in vigilance or complacency and how that changed in light of the true nature of the study. However, one participant replied specifically that, after finding out that there might be errors in the code, they are now questioning why they switched their answers when prompted. This participant was in the control group. The feedback provided to this participant was actually correct – all errors that the participant made were legitimate errors. It is possible that the participant did not read the full debriefing and was not aware of what group they were in and were simply trying to find justification for their mistakes.

One other participant from the deceived also answered the question about their thoughts about the process of using the simulation after finding out it contained errors in a very negative way. They specifically mentioned being upset that they switched their answers because the simulation told them to and yet the simulation itself was wrong. The surprise here isn't that a participant was upset, particularly one who was actively deceived and not warned, but that only one participant replied in this manner. Although this will be discussed later in relation to the deception in the study, this may suggest that this participant highly relied on the simulation.

After coding the questionnaire, the researcher also coded the interviews of the four participants willing to be interviewed. Only two of the interview questions had responses related to switching. The first, about how often the simulation disagreed with the provided answers, had the participant from the control group state that they often switched their answers because the simulation told them to. The second question, about what the participant did when the simulation disagreed with their answer, had two participants specifically say that when they disagreed with the simulation, they just switched their answers to make the simulation happy. One was from the deceived group and the other was from the misled group, both who had faulty simulations. This is not much of a surprise – the simulation is telling them their results are wrong – but it is still striking that, regardless of their knowledge, both of these participants still went along with the simulation despite expressing concerns about it. No matter their reason, this still shows a deference towards the simulation.

Factors. All the information explored in this question suggests that there is a connection between the correctness of the simulation and the amount of complacency shown by the participants. More complacency was shown by those with a faulty simulation, but all groups showed some level of complacency. Given that these participants are relatively new to the method of joints and it is their first time using the simulation, the Theory of Technology Dominance suggests that this should be the case.

Almost all participants, even those who were warned ahead of time about the possibility of errors, showed that they at least listened to the computer simulation when asked to check their answer and resubmit. This number aligns with the fact that only one participant got all fourteen forces correct initially and didn't change their response. Similarly, only four participants got all fourteen answers right after being given the chance to change their answer (including the original one). One of those who corrected their answers to the right answer was in the misled group – this appears to be the only participant who took the warning about errors seriously enough to not follow instructions by the simulation and to make changes that were appropriate. The other two with perfect final accuracy changed their answers based on correct feedback. The one misled student showed strong self-judgment and self-reaction in the face of negative feedback from the simulation and, as is shown by their final accuracy, the self-regulated learning appears to reinforce the correct method for the problems despite the simulation and stands to represent a key attribute worth refining in our students to help provide resilience in them when they are faced with misleading information.

Another common factor shown in the responses is a general distrust for the system, as shown by participant feedback based on their answers pre- and post-debrief in the questionnaire. Despite rating trust high, most participants stated that they were beginning to suspect something was wrong. This happened most amongst those in the deceived and misled groups, which makes perfect sense. These are the participants who are being compared to a faulty simulation – if they do the work correctly, they're going to be told they are wrong. What is interesting, as noted under switching above, is the one participant who was in the control group, got many answers incorrect, but chose to blame the simulation for their errors. This participant was only given correct advice and yet still chose to blame the technology and its "faults" for the errors that occurred. The fact that the distrust only really after the debriefing also makes sense, as even the informed and misled participants switched answers in ways that were expected. It also shows that many students allowed the simulation to replace or supersede their self-reflection phase during their learning.

Sadly, another common finding that emerges in participant responses is that the participants seem to show that any vigilance they learned by virtue of being a part of this study is directed squarely at this simulation. The researcher is interpreting the comments from the participants as being focused on this simulation itself rather than on simulations in general. It is as if finding out that this simulation was deliberately designed to test reliance means that they only have to worry about this situation when it is being tested for. It was hoped that participants would gain a greater desire to be more vigilant about using simulations and technology in the future based on this experience, which would help these participants be less likely to exhibit automation complacency in the future.

It appears that gender does not play a major role in whether students are more or less complacent based on gender. There is a minor tendency for younger students to accept that they are wrong when prompted by the simulation, but all age groups show a tendency towards complacency, suggesting that for the most part, there is no dependence on age to show automation complacency. Finally, it also doesn't appear that there are discrepancies based on a students' major.

Research Question #3b: What factors explain the impact that automation bias has on a participant's ability to recognize errors in that simulation?

As with subquestion 3a, the qualitative and quantitative information was triangulated to help answer this question. The insights found relative to the RW and WR switches helps to inform automation bias as well as automation complacency, which was addressed above. First, the quantitative responses to the Likert scale questions related to the Reuse Intention and Perceived Similarity found in the Reliance on Technology Questionnaire were evaluated to look for factors. Then, the qualitative coding of the free response questions found both before and after the debriefing during the Reliance on Technology Questionnaire were evaluated to look for factors relating to automation bias. Finally, the interviews of willing participants following the questionnaire were coded to look for these factors.

Reuse Intention. In looking over Table 36, Table 37, and Table 38, it is clear that participants with a correct simulation have a higher intention of reusing the simulation. Significant differences were found between those with a correct simulation and those without, and these differences were found for all five subquestions. These subquestions were whether the participant would use the guided simulation to help with the same solving task, to solve the same truss, to solve similar trusses, to aid in solving future problems, and to assist them with solving future problems. Differences were found based on treatment group, where participants in the control group were much more likely to reuse the simulation than those in the misled group. The subquestions for reuse intention show the exact same pattern, with those in the control group being significantly more likely than those in the misled group to reuse the simulation for the same task, a similar task, a similar truss, a future truss, or a willingness to reuse the simulation at all. This does not come as a surprise, as those in the control group had a working simulation with no cause to doubt it (they were not aware of potential errors), while those in the misled group had both a faulty simulation and were informed it might contain errors. The deceived group was not statistically significantly different than any other treatment group in Reuse Intention, which suggests that they have some reuse intention despite having a poor experience with the system.

This is evidence that it takes both knowing about possible errors and actually having them to influence someone against reusing technology. It also shows that when a participant's own self-judgment matches the feedback from the simulation during their self-reflection phase, the self-regulated learning cycle builds that participant's ability to perform correctly and their belief that the simulation was worth reusing.

Similar results were found in regards to the reuse intention and its subquestions related to trust, with the exception that no difference was found for reusing the program with a similar task. Those with high trust were significantly more likely to reuse the program again than those with low trust overall and for each subquestion except for reusing the program for a similar task. Again, there are no surprises, since those with high trust are showing faith that the simulation is working as intended, so we would expect them to be more likely to reuse the simulation.

Perceived Similarity. As pointed out in Chapter IV, there were no significant differences found between any groups for participant's perception of the similarity of the simulation's method for solving trusses and method they were taught. This is worth noting – there does not appear to be any groupings that found the process more or less similar than their own method. This is not to say that the method used by the simulation was similar to those used by the participants. Rather, regardless of the way the participants are divided, they responded the same way to the questions about process similarity. This may be due, in part, to the problem-solving process being built into the simulation very similar to the process taught in class. Those participants who used what they were taught in class found the process to be similar, while those who did not use the method they were taught in class would likely compare the simulation to their own process and may not find it to be as similar.

Free Response Qualitative Analysis. The coding of the free responses to the questionnaire questions focused on the written responses to questions about trust, help, and awareness. These codes were selected a priori based on the works of Billings et al., (1976) and Wandtner (2018). A response that showed that the participant trusted the program was marked with the code Trust. A response that showed that the participant was aware of the simulation's state, whether correct or incorrect, was marked as Awareness. A response that showed that the participant used the built-in help of the simulation was marked as Help. After comparing the results of the codes overall and based on the participant's simulation correctness, treatment group, self-reported confidence, trust, and experience (see Table 43, Table 53, Table 54, Table 55, Table 56, and Table 57) and based on their demographics (see Table 58, Table 59, and Table 60) the following patterns were found.

Trust. Prior to the debriefing, twenty-one (21) participants indicated that they trusted simulations to help them learn a topic. This is lower than was expected, as the Theory of Technology Dominance indicates that those new to a topic are more likely to trust technology. This may be due in part to the lack of experience with simulations shown by the participants – twenty-four (24) of the participants rated themselves as using technology to aid their learning as low. It may also be due to the experience of using the simulation and the knowledge some received that the simulation might be faulty impacted the response. Twelve (12) of the participants were using a working simulation and the other nine (9) were using a faulty simulation. In particular, eight from the control group, five from the deceived group, and four each from the informed and misled groups stated that they would trust technology. It is no surprise that the group who had a correct simulation and no knowledge of simulations would come out on the other side of the simulation and still have faith in technology. Each of the other groups had either a faulty simulation (the deceived and misled groups) or were forewarned of potential errors (the informed and misled groups). This likely explains the lower number who express trust in using guided simulations in learning a topic. Another surprise was that nineteen (19) of the participants who showed they had confidence in their work stated they trusted the system while only two (2) with low confidence stated the same. The surprising factor here is that only half of the low confidence participants stated that they had trust in the system.

Thirteen (13) of the participants who stated they had high trust for the simulation answered this question stating they had trust, which was another surprise. It was expected that more of those who self-disclosed having high trust would have been marked with a Trust code. Similarly, it is surprising that five (5) of those who self-disclosed as having low trust were coded with a Trust code. This discrepancy may be due to the interpretation of the questions by the participants. The self-reported trust was a single question that specifically asked about the current simulation they were using while the question asks about guided simulations in general. Some of those who don't trust this particular simulation may still believe that simulations as a whole are a good way to learn. They could have had a faulty simulation, leading to their lack of trust in this one, but they may believe that most simulations are still inherently positive in helping with learning. On the other hand, those who do trust this simulation, due to their experiences, may not trust simulations in general. This particular one worked for them, but they need proof from other simulations before they trust them.

This may occur due to a breakdown in the transition from the self-reflection to the forethought phase of the self-regulated learning cycle (Zimmerman & Moylan, 2009). If participants allow the feedback from the simulation to replace their self-reflection, their trust in the system may grow. If participants don't allow the simulation to replace their self-reflection, then the feedback may either bolster their self-reflection (when their answers agree) or it may hinder their self-reflection (when their answers disagree). When there is dissidence, it may change their feelings about simulations in general, impacting the way they answer this question.

Females tended to trust in simulations more than males, with 7 of 9 indicating trust in their answer to this question as opposed to only 14 of 25 males. The age of the participants also showed some differences, with all of those 25 and older trusting simulations while only 8 of 13 of those 22 to 24 and 10 of 18 of those 18 to 21 indicating the same trust. This suggests that older students are more trusting of simulations in general. Similarly, there were differences in trust based on major, with mechanical engineering students showing less trust (11 of 21) than civil engineering students (7 of 9) or other majors (3 of 4).

When asked how they checked their answer before resubmission, twenty-nine (29) participants stated that they did so because they trusted that the simulation was helping them work through the problem. This showed up most for those with a correct simulation, where all but one participant was coded as trusting the simulation, but it also showed up for those with a faulty simulation, where thirteen (13) of the seventeen (17) participants indicated trust in the simulation despite having the simulation comparing them with incorrect solutions. This correlates with the number of Wrong to Right (WR) Switches made by those in the control and informed groups and the number of Right to Wrong (RW) Switches made by those in the deceived and misled groups, where participants are making changes according to the prompt of the simulation and not the reality of the forces acting on the truss. It is also interesting to note that everyone from the control and deceived groups indicated they trusted the simulation as they rechecked their answers, while all but one from the informed group and all but three from the misled group indicated trust in the way they described checking their answers. This suggests that those who are not made aware of potential errors in the simulation are more likely to trust it than those who are made aware of those errors, regardless of if those issues exist. Within the group that were made aware of potential errors prior to their work in the simulation, those who had a correctly working simulation were more trusting than those who had a faulty simulation. It seems that it takes both being forewarned of potential errors and having a faulty system before most participants behave differently than the Theory of Technology Dominance indicates that they will.

Both males (21 of 25) and females (8 of 9) showed high trust when checking their answers. Similar to when this question was coded for complacency, older students showed more trust in the simulation than younger students, with all of those 25 and older and 12 of 13 of those 22 to 24 were coded for trust while only 14 of 18 of those 18 to 21 were coded for trust. This is all high, but there is a marked pattern. The participant's major did not have any major differences, with 17 of 21 mechanical majors, 8 of 9 civil majors, and all of the other majors coded for trust on this question.

When asked if the help prompt provided by the simulation was useful in solving the truss, only half of the participants (17) stated that they trusted the simulation enough to use its help. The majority (13) with this code came from those with the correctly working simulation. Surprisingly, this came from the informed group, where every participant said something to the effect of trusting the simulation and the help it provided. Only five (5) of those from the control group did the same. It appears that knowing about the potential errors and not encountering them led the informed group to be more trusting of the help provided, while those who had no indication there could potential errors were less likely to trust the system. On the other hand, of the four (4) participants with faulty simulations whom were coded for trust on this question, three (3) came from the deceived group and only one (1) came from the misled group. This aligns with those with a correct simulation – those in the misled group were informed of potential errors and were therefore more likely to spot the errors that occurred, while those in the deceived group may have just felt that they were wrong, showing that they did, in fact, trust the faulty simulation with its inherent errors. Their self-reflection favored the results of the simulation rather than their prior experience. It is also striking that sixteen (16) of those who showed trust when using the help feature reported having high confidence in their own skills while only one (1) who showed trust reported having low confidence. It was expected that more of those with low confidence would have had shown more trust in the system, but despite their lack of confidence, they may have had poor interactions with the feedback from the simulation and begun to doubt the simulation during the self-reflection phase. By being less confident in their ability, they may have been more questioning of the state of the simulation and therefore had less trust.

The last question related to trust in the pre-debrief portion of the questionnaire was about what the participant did if they disagreed with the simulation. Based on the previous responses, the researcher was surprised to find that only four participants, one from each treatment group, indicated that they trusted the simulation when they disagreed with it. The wording here may have helped force this result, as asking about disagreements may have caused further self-reflection for the participants before they answered and may have given some participants a glimpse of the actual purpose of the study. It should also be noted that all four participants who were coded for trust on this question had self-reported high confidence and high trust.

There were two questions that were coded for trust that had results surprising to the researcher. The first of these, when asked if participants changed their answer when given the opportunity and why they did so, had five (5) participants specifically say something coded as trust for the simulation. Two (2) participants came from the control group, while each of the other groups had one participant coded this way. This question was primarily aimed at automation complacency and, specifically, whether the participant switched their answer, but when answering the why portion, these five all said something to the effect of "because the simulation told me to," implying a level of trust in the simulation and ceding any questioning level of self-reflection to the simulation.

The other question with a surprise result in the coding was on the question of whether the simulation ever told the participant that their answer was wrong and whether they agreed with the simulation. Three participants stated something that was coded as trust. Interestingly, all three had high confidence, suggesting that they felt their answers were right prior to being checked by the simulation. Despite this, they all claimed that they changed their answer specifically because it was what the simulation wanted. Two of these participants self-reported having high trust, while the other was neutral, so this coding is in line with their self-reported trust of the simulation. Two of the participants came from the deceived group while one came from the informed group. It should be noted that none of the control group responses were coded with trust. This may be due to them having a correctly working simulation and getting correct feedback such that overtime, they had less disagreements with the simulation. Thus, they may have had been coded for trust had the question been phrased differently.

Following the debriefing, the participants were asked further questions related to trust. When asked how the participants felt about not knowing the purpose of the simulation from the beginning, only one participant mentioned anything related to trust. Their specific comment actually shows a trust in simulations as a whole, and they stated they felt betrayed because they should have been able to trust the simulation. Most reported on the nature of the study itself rather than on anything related to their trust in the system.

When asked if participants suspected there was something wrong with the simulation, eleven (11) participants were coded for trust. Unsurprisingly, five (5) of these came from the control group and three (3) from the informed group. Both of these groups had working simulations, so their self-reflection phase allowed them to build correct mental models which, when performed in future problems would lead to more correct answers. When they made mistakes, they would be able to self-reflect and correct the mistakes they were making. What is surprising is the two (2) from the deceived group and the one (1) from the misled group that stated they didn't suspect anything was wrong with the simulation. All three (3) of these participants had high confidence in their answers, so they either changed their mental models while self-reflecting on each answer to an incorrect model after the first few problems presented by the simulation or they were already had an incorrect model. Otherwise, they should have had lower confidence in their own work.

When asked about how the intentional errors impact their thoughts on the simulation, sixteen (16) of the participants were coded with trust. While most of these responses (10) came from those with a correct simulation, there were still six (6) who came from a faulty simulation. Most of these responses indicate trust by inverting the comments, stating that they should have trusted themselves more than the simulation. This suggests that these students were replacing their own self-reflection in favor of the simulation, a further sign of the automation bias being present. Even after the debriefing, these sixteen (16) participants were willing to recognize that they did put their trust in the simulation.

When asked if the knowledge of the study changed their perception of the help provided by the simulation, only five (5) participants were coded as showing trust in the simulation. Four (4) of these participants had a correct simulation while one (1) had a faulty simulation. Those with the correct simulation were split equally with two (2) participants from both the control and informed groups. All four with the correct simulation stated that their views didn't change from before, which makes sense since they had no reason to doubt the simulation. The one participant with the faulty simulation came from the misled group and, after being informed that they did have a faulty simulation, indicated that they now question the value of listening to this simulation; i.e., they have indicated that they trusted the faulty simulation from the beginning.. They trusted before and only question after being informed of the truth. It is interesting that the one from the faulty simulations that indicated trust came from the misled group. This participant should have been more wary of the simulation, especially as it continued to "correct" the participant's work.

When asked about their thoughts on the process used by the simulation now that they know about the intentional errors, nineteen (19) were coded as showing trust in the simulation. Most of these nineteen codes are similar to the above – expressing that they should have trusted themselves while working through the simulation and, thus, that they were trusting the simulation at the time. There were six (6) participants from the deceived group and five (5) from the misled group that were coded this way. The three (3) from the control group and the five (5) from the informed group expressed a different viewpoint, being more grateful that their trust in the simulation was rewarded in that they did trust it, but it was worth trusting. This does suggest that, prior to being told, some participants did trust the system inherently even though it provided incorrect information to them during the self-reflection phase. This is further backed by the fact that nine (9) of these participants showed high trust, seven (7) of them with correct simulations, while eight (8) of these participants showed low trust, all of which had a faulty simulation. None of the participants who were coded for trust showed low trust. There were two (2) participants with a faulty simulation, both of which were in the misled group, which reported having high trust in the simulation.

When asked how this study impacted the participant's perceived usage of simulations moving forward, only eight (8) participants indicated that they trusted simulations moving forward. A few recognized the situational nature, where they would trust a simulation that they knew was correct, but three (3) indicated that they use enough technology that they inherently trust that technology until proven otherwise. This suggests that most participants are willing to use technology, but only small portion consider ahead of time whether they should rely on it or not. It also suggests that, during the self-reflection phase, most students may wonder whether the state of the simulation is correct but, ultimately, they still let the feedback from the simulation impact their self-judgment.

The interviews of willing participants following the questionnaires were also coded for trust. A specific example involves two interviewees from the misled group who were asked how they used the method of joints to solve the provided truss, they both mentioned something about trusting the instructions provided by their simulation. What is particularly interesting is that both participants also rated themselves as having low trust in the simulation overall. Neither the one from the control group nor the deceived group specifically mentioned the simulation at this point. The one from the control group did specifically mention that the system helped him solve the provided truss problems, including the help sections that helped provide formulas that needed to be solved. Two interviewees the one from the deceived group and one from the misled group, also both indicated that they trusted the simulation when it said their answers were wrong at first, but over time, they started to doubt it more and more. So they started from a position of trust but that trust was eventually eroded. The interviewee from the control group and the one from the deceived group also both specifically mentioned that they trusted the system when asked how they reacted upon finding out the true nature of the study, although the one from the control group did mention relief that maybe their simulation was incorrect. This seemed like an attempt to justify why the simulation tried to correct some of their solutions.

Awareness. There were fewer questions in the questionnaire which were created for awareness. The state of the technology is defined as whether it is functioning correctly or not and lines up with whether the provided simulation is correctly guiding students to a solution or not. Awareness and vigilance are similar, but where vigilance is about actively checking the state of the technology, awareness is about expressing an understanding of the state of the technology. One can be aware and not vigilant – that is, they may correctly be aware that their technology is working correctly and therefore aren't vigilant about it. This is most likely to occur when the simulation is working correctly. If awareness is correct for an incorrectly working simulation, then vigilance should rise as the aware participant works with the simulation more.

Before debriefing, only one question was aimed at gathering participants' awareness of the state of the simulation. When asked whether the simulation told them their answer was incorrect, only twelve participants were coded as showing awareness. There were two (2) each for the control and informed groups, of which both groups had a correctly working simulation. While it was expected that more would have been more aware of the correct state of the simulation, it makes sense that, given a correct simulation, most participants might not even think about the state of the simulation and so wouldn't provide comments that led to the idea of awareness. While it is likely appropriate to assume that the others in these groups might also have been aware, the researcher determined that he had to limit the application of the code to just those who actively stated something about the state of the simulation.

Half of the deceived (4) and the misled (4) groups all had statements that indicated they thought something was wrong with the simulation when they answered this question. This is also expected, as they were working actively against a faulty simulation. Put another way, if their answers were correct, they would still be constantly questioned by the simulation and would thus become more aware of the state of the simulation as they went. Similar to those with a correct simulation, it might also be appropriate to assume that most participants had some awareness, but they aren't displaying any such awareness here and therefore their data was not coded for awareness. It's also possible, particularly for the deceived group, that participants actually felt like the simulation was in a correct state and in doing so, during the forethought phase, changed their plans on how they would approach future problems to the incorrect method.

Following the debriefing, no questions were intended to gather information about awareness. However, two questions other questions had one or two participants who indicated their awareness of the simulation in them. Both questions were intended to gather information about how vigilant the participant was, so it is not that surprising that a few participants' data could also have been coded for awareness.

When asked if the participants suspected whether there was anything wrong with the simulation, one participant from the deceived group and one from the misled group both stated that they suspected something as soon as the signs to the correct answers was wrong. It is not surprising that both of these participants came from the groups with faulty simulations – those are the participants most likely to notice that their work was being marked wrong incorrectly.

The fact that these two both specifically mentioned the actual error that was present is why they were coded for awareness, even though others from their groups mentioned they had grown at least somewhat suspicious of the simulation.

When asked about the process of using the simulation once they found out whether they had a correct or incorrect simulation, one participant from the misled group specifically stated that their process was different from what the simulation expected. They were the only participant to specifically reference what the simulation was expecting, showing at least some level of awareness.

Each of these surprises came from those who had a faulty simulation. This suggests that those who expressed awareness were most likely to come from groups that had to deal with the faulty simulations. This suggests that when a simulation either works correctly or aligns with the method the participant is using, that they will not be consciously aware of the state of the technology. It also indicates that the default assumption is that the state of the technology is right.

During the interviews, one interviewee from the misled group specifically mentioned that they felt that there was something wrong with the simulation as they described their experience using other technology and when learning the method of joints. The other three didn't indicate anything when asked about their use of technology, either generally or with the method of joints specifically. The same interviewee also mentioned their awareness when asked what they did when they disagreed with the simulation. Additionally, both of those in the misled group did mention that they were aware of having faulty simulations early on while working through the program. Finally, all four participants expressed that they knew something was up when asked how they felt when they found out their simulation might have contained errors. This final point is important to recognize – this aligns with the coding from vigilance, where more vigilant statements were made following the debriefing. This may relate to the participants finally putting words to what they felt, but regardless of how aware the participants are about the system, they still followed the directions of the simulation as shown by the number of appropriate switches made by the participants (either Right to Wrong (RW) or Wrong to Right (WR) based on the correctness of the simulation).

Help. The biggest surprises in the coding came when looking for whether the help was used by the participants. It was expected that many students, particularly those who had a faulty simulation, would use the help as they solved for the forces in the joints provided by the simulation. Participants accessed the help by clicking on a button labeled "Help" within the simulation. The first time they clicked on the button, it provided a general outline of how to derive the necessary equations. The second time they clicked on the button, it provided the equations to the participant based on whether the simulation was giving correct advice or not. However, as shown in the coding tables, the helps were not used as much as expected. Most of the questions related to help were asked prior to the debriefing.

When asked how participants felt about using simulations to help learn a topic, two (2) participants, both from the misled group, indicated that they thought that simulations could help students learn topics. One participant specified that they often use YouTube videos to help them better understand the material. Nevertheless, their numbers was much lower than expected, as one of the assumptions of the researcher was that students, in general, tend to use technology to better understand material. This could be somewhat related to the question itself not stimulating participants to responses that were coded to help – the question likely needs to be thought of more deeply if it is used in the future.

When asked if they used the help, twenty-one (21) participants mentioned that they did. This shows that during the performance phase, participants engage in help-seeking. However, twelve (12) participants mentioned that they only used it one, two, or three times and generally found it to be unhelpful. Thus, during the performance phase after trying to use it a handful of times, these participants did not get the support from the help they needed and so stopped using it. Five (5) participants from the control group and seven (7) participants from the informed group stated that they used the help, while six (6) participants from the deceived group and three (3) from the misled group stated they used the help. Those with a correctly working simulation were more likely to use the help, which is interesting since those with a faulty simulation were expected to need the help more often as a correct answer would be marked wrong. This may be because they were better able to apply the self-reflection phase and, when the help was used, it aligned more often with their own analysis of how to use the Method of Joints. As such, the self-regulation cycle may have been followed more closely, leading them to use the help more often in the future (Zimmerman & Moylan, 2009). Those with a faulty simulation may have fought with the help more, deciding that it wasn't helpful to them, and as such, the self-regulation cycle may have dragged them further away from the help rather than reinforcing its use throughout the simulation. Three (3) participants specifically stated that the help didn't actually help – it was either confusing or just lacked the details the participant was looking for. This is probably the most disappointing result for the researcher, as it means the work done to help the students failed. It is possible that some of the negative comments came from participants who weren't aware that there was further help or what that extra help entailed. Only one participant responded to this question in a way that implied they accessed the second level of help.

The demographics based on help are also different. Only 14 of 25 males indicated that they used the help while 7 of 9 females made the same indication. It appears that females used the help more often than males. Those 24 and younger also indicated they were more likely to use help, with 11 of 18 of those 18 to 21 and 9 of 13 of those 22 to 24 indicating they used the help, while only 1 of 3 of those 25 and older indicated the same. This could be due to the low number of students older than 24, but it is still worth noting now. The participant's major did not appear to make any impact on the whether the participant used the help, with 14 of 21 mechanical, 5 of 9 civil, and 2 of 4 of the other majors coded for help.

The following question asked participants how effective the help was. Nineteen (19) of the participants mentioned that the help was useful, showing that they used the help during the performance phase. This is interesting since so many mentioned they only used it a handful of times and thirteen (13) indicated they didn't use the help at all. Two (2) participants were coded for help on this question that weren't coded for help on the previous question. Four (4) participants who were coded for help on the previous question were not coded for help on this one, including all three (3) of those with negative responses about the help. Six (6) participants came from the control group while all eight (8) from the informed group stated that the help was useful. Four (4) of the participants from the deceived group and one (1) from the misled group also mentioned that they found the help useful. Eighteen (18) of these participants had high confidence in their answers originally. This is supported by the comments – sixteen (16) mentioned that it helped them make sure the equations were right. Two (2) of those who made this comment came from the deceived group, meaning they accepted the faultiness of the simulation and started to solve the problems that way. One participant from the informed group specifically mentioned that the help actually made the work too easy – it provided all of the formulas, so it was trivial for that participant to solve each problem provided. The fact this participant was from the informed group is significant – although it is not known for sure, the fact that they were warned of potential errors makes their use of the help more important.

There were two questions from before the debriefing for which participants were coded for help that weren't expected. The first code, how participants checked their answers prior to resubmission, had one participant from both the control and deceived groups specifically mention that they relied on the help as they checked their resubmission. The second, asking what participants did if they disagreed with the simulation, had a different two participants, one from each of the control and deceived groups. Both of these participants sought to figure out why the simulation told them their answer was wrong and both indicated that they used the help button to help them figure out where they went wrong. The most significant part of this is that neither of these participants were informed of potential errors. These participants were the most likely to trust the simulation from the beginning of their work.

Following the debriefing, there was only one question specifically designed to address the help provided by the simulation. When asked how their perception about the help changed when they found out the true nature of the study and the simulation, six (6) participants specifically referenced the help in their answers. Two came from each of the control and the informed group, while one came from each of the deceived and misled groups. Both of those with the faulty simulation specified that they now question the help, but both used it previously and found it to be helpful. So, they had a negative change in perception, which is to be expected. The four (4) with correct simulations stated that the help still seemed useful to them, which is a good sign since they had correct help provided. One (1) of them did note how easy it would be to deceive a user with incorrect ratios or signs. These results suggest that the correctness of the help is directly related to whether it will be viewed as useful, which is what was expected. This also ties the help-seeking performed by the participants during the performance phase and the self-judgment that occurs during the self-reflection phase.

Like the questions before the debriefing, two questions given post-debriefing found participants responses coded for “help” on questions not really intended to find that. The first, about how the study impacts the participant’s perceived usage of simulations moving forward, had one (1) participant from the informed group mention that they still believed that using simulations was helpful in learning about the methods. The second, about how the participant would ensure future uses of technology wasn’t faulty, had five (5) participants, two (2) from both the deceived and informed groups and one (1) from the misled group, each mention that they would seek help from alternative sources to verify the results of a simulation. Although neither is quite the same as the help provided by the simulation, it does suggest seeking help from a reliable source in using simulations. No one from the control group provided statements that could be coded for help, suggesting that those with a correctly working simulation, and no reason to doubt, it were less likely to be critical of either the help or other simulations directly.

The interviews seemed to mirror the first big question about help – only two of the interviewees even used the help, with one originating from the control group and the other from the deceived group. The one from the control group also specifically mentioned using the help in their strategy for solving the trusses. One of the participants from the misled group specifically mentioned that the one time they looked at the help they found it to be confusing and not worth using.

Factors. All the information gathered about automation bias has been evaluated and compiled, and it has shown a few factors that lead to automation bias. The fact that differences were found between those with a correctly working simulation and those with a faulty simulation was found to be significant very often. Regardless of the simulation style, all participants demonstrated their trust. This is shown by the number of Right to Wrong switches made by participants with a faulty simulation and Wrong to Right switches made by participants with a correct simulation. In addition, prior to debriefing, very few students indicated an awareness of the state of the simulation. This shows that, regardless of the state of the simulation, the participants weren't questioning the correctness of the simulation while using it. In addition, the desire to reuse this or similar simulations was higher for those with a correctly working simulation, as well. Those with the correct simulation receiving correct advice are the most likely to gain in confidence as they work, which should help them determine that they want to continue using similar simulations. Those who constantly had to fight with the simulation, being told they were wrong at many steps, are less likely to want to use a program that tells them they are wrong. Of course, this wasn't universal, as there were some participants who fought with the correct simulation and some participants who agreed with the faulty simulation. However, the fact that the willingness to reuse the simulation was significant between the two groups is of significant note. In digging further, it appears that the group least likely to reuse the simulation are those who were misled about the study – they were warned ahead of time that there might be potential errors and they constantly fought with the simulation, leading them to believe that it this type of simulation isn't helpful or useful.

One factor leading to automation bias in participants is that students often believe that simulations provided to them in a class are intended to help them learn. That is, they arrive in class with a preconceived notion that any simulations used are correct. Many participants showed trust in the simulation prior to the debriefing, showing that they were following its directions during the performance phase and making changes based on its feedback during the self-reflection and forethought phases. Others were willing to use the help, especially at first, to determine how to use the simulation. One participant even specifically mentioned that the deepest level of help was too revealing, as it allowed them to find all the values without having to think about the formulas. There were also those who, following the debriefing, stated that they were disappointed that they had trusted the simulation and that they should have listened to themselves. This also lines up with both the decision switching that occurred, with a higher percentage of decisions switched from wrong to right answers for those in the deceived and misled groups as well as being reflected in the initial and final accuracies. This part of the analysis indicates that these participants are expecting simulations to help them learn. It helps that, for most of these students, this was their first time learning about the method of joints – that is, these participants were part of the naïve newcomer crowd that is more likely to have automation bias.

Another factor of interest involved the skill level that the participants had coming into the study. The average correct number of forces found was 56%, suggesting that even before receiving feedback, many participants were struggling with the concept. The instructor does mention that this material takes some practice for students' confidence to become high. One participant had 100% accuracy both before and after receiving feedback, and they happened to be from the informed group. They were joined by three more participants who happened to end with 100% accuracy. One of these participants was actually from the misled group – they originally had a 64% accuracy, but after being told their answer was incorrect, they found the correct way to work through each joint and correctly solve the forces present despite the help provided to them being incorrect. Every participant who had a correctly working simulation had either no change in or an increase in accuracy before and after receiving feedback, suggesting that the feedback helped them find the correct answer. The participants with a faulty simulation had a more mixed response, with some showing an increase in score and others showing a decrease in score as they were misled by the simulation.

The context of the problems also seems to be a factor in how the participants respond to the simulation. Many participants pointed out that they checked their answers before resubmitting them because the simulation asked them to. Some of these expressed regret during the debrief that they hadn't trusted themselves more, showing that they might have had some awareness of the state of the simulation, but not enough that they trusted themselves over the simulation. All statements indicating the participant knew something was wrong following the debriefing may help the participant feel better about themselves, but it also indicates that they have automation bias – they specifically trusted the technology over themselves despite evidence that they shouldn't. It's important to note that this behavior appeared for members of all treatment groups, including the control group. It appears that some participants are looking for many possible reasons to excuse themselves from responsibility for their choices.

The help provided to the students ended up being a factor in the level of trust provided to the participants. Almost all participants stated that they used the help at least once, but many stated they only used the help a few times. In addition, although a few mentioned that they benefitted from the help, many more found it to be distracting at best and a hindrance at worst. Originally, it was unclear how many students actually used the second level of help, as only a few mentioned getting the equations from the help. Those who did use the equations, whether they were correct or not, showed that they had high trust in the simulation regardless of what treatment group they were in. Those who found the help to be less useful had lower trust, as well.

Thus, regardless of the correctness of the simulation, the level of help used by the participant appears to be correlated with the trust that participant gives to the simulation.

Gender appears to be a factor for participants, with females more likely to trust simulation in general and being more willing to use the provided helps found in the simulation. Age also plays a factor, with younger participants using the help more often than older participants but older students being more likely to trust the simulation and check their answers when the simulation informs them their work is wrong. Finally, it appears that mechanical engineering students are less likely to trust simulations in general than those from other engineering disciplines.

These factors suggest that most of the participants using the simulation had high automation bias, trusting the simulation despite any evidence that they shouldn't. There was very little awareness shown by participants in any treatment group, showing that participants were not actively thinking about the state of the simulation and whether it was correct or not as they worked, with very few exceptions. This automation bias can be explained by the Theory of Technology Dominance. These naïve newcomers to the Method of Joints were also new to this specific simulation and to guided simulations in general, as shown by their responses to previously used educational technology. The Theory of Technology Dominance suggests that people in this exact situation will show high trust in the simulation, and, in general, that is what was found here. This is significant because of the vulnerability of a novice learner and may suggest that faculty always vet educational simulations well before implementation.

Use of Deception

This section will discuss the use of deception in answering the research questions. First, the deception was necessary in order to gather the data related to both automation complacency and automation bias. Awareness was low throughout the study, as even following the debriefing experience, the vocabulary that coded to awareness rarely appeared. The participants didn't really reflect on the state of their simulation well even after being informed of the state, especially for those who were given a correct simulation. Vigilance increased after the debriefing, suggesting that if the participants had been informed ahead of time, they would have been more vigilant. In addition, the switching that occurred and the initial and final accuracy of answers submitted were based more on the correctness of the simulation rather than on the participants understanding the process of solving using the method of joints. It is likely that these results would not have been the same if participants had known the true nature of the study – they would have either shown that they had the bias or they would have ignored the simulation entirely, ignoring both the feedback and the help provided. The deception was necessary to find the answers to the research questions.

Deception studies require additional care both prior to and following the study. Moral turpitude, the question of if the deception can be considered morally acceptable, is the first consideration. It is important that the information gained outweighs the risks to the participants. The risks and the mitigating circumstances related to them will be discussed below, but it is important to note that this study showed that automation complacency and automation bias were present amongst the participants. As such, this study falls in line with others that have come before, showing that the Theory of Technology Dominance applies within the field of Engineering Education just as it does within Medicine (Goddard et al., 2012, 2014; Kemmelmeier et al., 2003; Uz & Kemmelmeier, 2017). Although it could have been assumed, having this evidence can inform engineering education professionals in teaching, instructing, and understanding their students. Thus, this level of deception was necessary at this point.

The rights of the participants were also upheld during the course of the study. Seventeen (17) participants withdrew from the study by never completing the post-questionnaire and eighteen (18) immediately upon being debriefed. No pressure was used to try to encourage these participants to stay in the study – they had the right to withdraw, and this was honored. In addition, no information about the participants was viewed as harmful, embarrassing, or private to the participants. Their anonymity was secured and it would be incredibly difficult for any individual to be identified.

Although there was potential harm to the participants in potentially learning an incorrect process for the method of joints, the responses from the participants of all groups indicates that there is little reason to worry. The biggest group that may have received harm are those who withdrew at any point, particularly those before the debriefing, but it is possible that those who did not complete the post-questionnaire either did not start using the simulation or became so frustrated with it that they didn't use it long enough to learn the incorrect methodology. In addition, the method of joints was not one of the topics on the final exam, so the participants' grades were not likely to be negatively impacted by an incorrect method. What is not clear, however, is how students who used the faulty simulation ensured that they were familiar with the proper process for the method of joints – there was no indication that any participant used the simulation again to learn the proper process.

Potential harm to the discipline is also likely not going to be very high. There were no comments in the Reliance on Technology post-questionnaire, pre- or post-debriefing that indicated that any participant viewed this study as negative towards their engineering disciplines. When asked when the appropriate time to know that a simulation is faulty, most stated that they want to know at the beginning. While this simulation intentionally had errors, it is possible that other simulations will contain unintentional errors, which is hard to warn students about. Some of these comments even came from those who were informed of possible errors, meaning that even when forewarned of errors, many of these participants still started expecting the simulation to be working correctly and to be an appropriate way to learn the method of joints. One potential harm of this study to the discipline is that it may reinforce the participant's understanding of where errors come from in technology – they may assume that all errors are intentional rather than unintentional.

Participants responded to how they felt when the deception was revealed. Although this question was coded for the Trust they showed in the simulation, it is also important to analyze their answers here regarding the deception. The most common response was from fifteen (15) participants who found the deception necessary. Thirteen (13) were neutral or had no comment about the deception. Three (3) mentioned that they were upset during the process and now know why. Finally, three (3) said that they found the experience to be amusing or even good. What's most interesting is that of the three (3) that were upset during the process, two (2) were from the informed group. Both expressed that they were annoyed that the program was constantly correcting them. When informed of the deception, it appears that both of these participants seem to believe that they were in the deceived group, even though they both had a correctly functioning simulation. It may be that both took the notice of potential errors to heart and assumed that their simulation had errors, even though it really didn't. The other participant, who stated that the deception made them angry while using the simulation, was from the deceived group. No other participant from the deceived or misled group expressed anything more than being ok with the deception or upset about the process. With only three participants expressing any frustration over the deception and that frustration being during their work in the simulation, it appears that the participants accept the deception as a necessary part of the study. This is supported by all four interviewees stating that they thought the deception was rather well done. One even stated that they laughed for ten minutes upon being told about the deception and that they then felt a lot better about their original work. None of the four felt that the deception shouldn't have been a part of the study.

It is also important to point out that all participants had ample opportunity to recover from learning incorrect techniques. All students were provided with a code for a correctly working simulation, and both the instructors and teaching assistants were available to reaffirm correct solution process and procedures.

Implications

The results of this study have shown that there are several implications related to both the use and reliance of technology and the use of deception studies. First, students have both automation complacency and automation bias. They do what they are asked to do, even if they have misgivings about how they are using that technology. Participants mentioned that they felt something was off about the simulation, and yet they still switched their answers and had final accuracy in accordance with the correctness of their simulations. For those with a faulty simulation, this is contrary to the method they were taught in class – they should have been questioning the simulation. This was also true for those who were informed of potential errors. The responses from these participants imply that they either ignore the notice, starting with a large amount of trust in the simulation, or they use the fact that there are potential errors to justify why they are disagreeing with the simulation.

Students give a lot of trust to technology, and this trust may be even further exacerbated when that technology comes from their instructor. It is impossible to expect that all technology will be free from errors, so faculty should be familiar with the technology they choose for the students to use. In addition, faculty should put emphasis on ways to ensure that the technology is being used correctly. This may involve ways to build familiarity with the results of the technology. It may also provide justification for why faculty require students to work problems by hand. Sometimes knowing the process more intimately can make the errors in the technology more noticeable.

Students who experience the negative consequences of faulty technology still show that they seem to rely on technology more than they need to. When asked about their use of future simulations and when they should know about potential errors, participants stated that they should know before using the technology. This response even came from those in the informed and misled groups, who were informed ahead of time. The conversation of automation bias and automation complacency needs to happen often to help the students learn how to combat these tendencies. It is unlikely that only one time will work for most students. Although twenty-six participants stated that they would check future simulations against other sources, none of them did while working on this simulation. It is unclear if any of these participants will check future simulations, although it can be hoped. Having more experience with faulty technology and having more discussions about recognizing faulty technology will likely help students better prepare for being vigilant when using technology in the future. This should help combat the automation bias and automation complacency that tends to arrive.

Finally, faculty should be aware of the Theory of Technology Dominance and how it might apply to them. Even a seasoned faculty member in a close field to engineering who tested the simulation for the researcher fell for the errors in the simulation. He had followed the instructions carefully, questioned them a little, but ultimately decided that since it was designed by the researcher, he could follow its instruction. He reported later that he figured his memory might be off, and he switched his solving strategy to match the faulty simulation. It wasn't until he was debriefed by the researcher that he realized what had happened. Faculty may also give this inherent trust to technology used in their course, such as that provided by a publisher or third-party software developer. By consciously being aware of the potential for errors and demonstrating to the students how an expert verifies the correctness of their technology, it is possible that faculty may be able to change the direct impact of the Theory of Technology Dominance on their students.

Recommendations for Future Studies

There are several ways in which this research can be further developed. There is an open question as to why thirty-five (35) participants failed to complete the study. In particular, why did eighteen (18) participants start the Reliance on Technology questionnaire and yet opt out after being debriefed? Because of the deceptive nature of the study and how participants were allowed to opt out during the study, this is an open question. It could be that they were the participants most upset about the deception, as eleven (18) of the eighteen (18) participants were given a faulty simulation. However, it is also likely that participants chose an easy out. They were assured that they would receive full extra credit without further work and they may have chosen to take advantage of that opportunity, especially since they knew there were further questions.

One way to address this would be to change the Opt Out options presented during the debriefing. Instead of limiting the options to simply staying in the study or exiting the study, participants could be given a third option of exiting the study but allowing their current data to still be used in the study. In this way, those who are upset about the study and want to leave are allowed to do so while those who just don't want to do more work but are comfortable with their data being used can leave and their data can remain in the study.

This tactic still would not account for the seventeen (17) participants who simply didn't start the final questionnaire. One recommendation to at least address this would be to build in to the study methodology a follow-up with any participant who starts the simulation. Many are busy and may not take the time while others may have given up using the simulation due to issues inherent in it. Having some additional contact between the researcher and the participants may help them continue in the study or give them a better opportunity to address why they are not completing the study. This could be as simple as messages between the researcher and the participants prior to the study ending, but it could also involve a conversation between the researcher and the participants. Any information that could be gathered from these participants can shed further light on the reliance that students have on their technology.

While it was expected that the interviews would be useful in identifying factors, most of the factors showed up in the Reliance on Technology questionnaire. This is partly due to the low number of participants who volunteered to participate in the interviews, but is also due to the fact that only four of the eight invited interviewees was actually interviewed. It may be better to randomly select interviewees from the different treatment groups. It may also help to provide further incentives to participate in the interviews – they are an additional amount of time beyond the work that other participants are putting in.

The Reliance on Technology questionnaire did help identify factors, but the qualitative questions could be polished for better understanding and alignment. The number of unanticipated codes assigned to different responses shows that the questions may not be interpreted by the participants in the same way that the researcher did. By further refining the questions, it may be easier to get the expected codes with fewer surprises while also not leading participants to their answers.

The overall results of the simulation show that these participants have high levels of automation bias and automation complacency, but there were only thirty-four (34) participants that were analyzed. This research needs to be continued in similar and diverse courses to help verify if the results can be generalized in full. Power analysis suggests that 180 participants are needed to generalize the results. Further research should continue the four treatment groups of control, deceived, informed, and misled to keep the studies similar enough to build up the sample size.

Other forms of simulations and technology also need to be used in similar studies. It is possible that it was the particulars, such as the interface and presentation, of how this simulation was built that lead to the factors identified by the researcher. By varying the technologies used, along with deception, it may be possible to identify those factors that are universal and those that may only apply to this group of participants.

Finally, research into other topics within Engineering Education should consider using deception to help find those factors that cannot be otherwise understood if the participants know the true nature of the study. This could be further researched in the reliance on technology field, but it may also be in topics such as the impact of the gender, gender identity, or race of a faculty member on student perceptions of learning. Carefully analyzing the purpose behind the deception and ensuring that little to no harm occurs for the participants, the researcher, the discipline, or society is important. When done correctly, information can be gathered and participants can still benefit from the topic being used in the research.

It would also be fascinating to deliver a similar simulation to an upper level course in engineering to see if results parallel those found herein. It is likely that cognitive maturity found in seniors as well as deeper engineering problem solving experience will change inherent automation complacency and bias. In a similar manner, simulation deception studies delivered in other disciplines and at other maturity levels beyond those found in the university could also be very interesting. Indeed, developing an understanding of how these two factors range in professional engineers may especially be important.

Limitations of the Research Design

This research was based on several different factors. The prequestionnaire was a typical demographic survey, although it included a few specific additional questions. It is clear that the wording of the question for previous experience led students to include all sorts of technology, ranging from learning management systems to online homework systems, most of which are not in the same category of computer simulations like the one created for this study. The postquestionnaire consisted of a modified version of an already existing and validated survey combined with new questions created specifically for this study. Although the new portions of this questionnaire were face validated, there appears to still be issues, as many questions had low levels of the intended coding while some questions received additional codes emerge during the coding process. In addition, there were a low number of participants in the study, meaning that the knowledge gained from these questionnaires may be transferable but is not yet generalizable to the larger population of engineering education majors. The interview protocol is similar. Although the questions were better aligned with the a priori codes after interviewing participants, there were far too few participants to generalize from the findings here. This is reflected in the minimal impact the interviews had when answering the third research question.

The simulation should also be considered as a limitation in this study. The simulation was coded specifically for this study. This means that it was not an already existing tool used by the participants. There could have potentially been errors, such as portions of the program not working as intended. The researcher coded the simulation directly and had two professors who teach the method of joints validate the software, including the main advisor of the research, and made changes based on this feedback, but it is possible that other errors may have been missed. In addition, the simulation does not necessarily look the most professional. The researcher utilized the default C# classes and forms to create the simulation. This may have impacted the participant experience using the simulation, including whether participants found the Help button easily. If students had a hard time finding the button, this may have impacted how much help they sought from the simulation. The text portion of the simulation was not adjustable, either, meaning that all participants had to use the same font size. This may have impacted their willingness and even ability to read the material provided. The quality of the simulation could definitely have an impact on the outcomes of the study.

REFERENCES

Agnew, D. M., & Shinn, G. C. (1990). Effects Of Simulation On Cognitive Achievement In Agriculture Mechanics. *Journal of Agricultural Education*, *31*(2), 12–16. https://doi.org/10.5032/jae.1990.02012

Alchemer. (2021, June 7). *How to Write Better Demographic Survey Questions (With Examples)*. Https://Www.Alchemer.Com/Resources/Blog/How-to-Write-Better-Demographic-Questions/.

Al-Natour, S., Benbasat, I., & Cenfetelli, R. T. (2008). The effects of process and outcome similarity on users' evaluations of decision aids. *Decision Sciences*, *39*(2), 175–211. https://doi.org/10.1111/j.1540-5915.2008.00189.x

American Psychological Association. (2017). American Psychological Association. Ethical Principles of Psychologists and Code of Conduct. *American Psychologist*, *57*(12), 1–20.

Arnold, V., & Sutton, S. G. (1998). The theory of technology dominance: Understanding the impact of intelligent decision aids on decision maker's judgments. *Advances in Accounting Behavioral Research*, *1*(3), 175–194.

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology*, *96*(3), 523–535. https://doi.org/10.1037/0022-0663.96.3.523

Azevedo, R., Winters, F. I., & Moos, D. C. (2016). Can Students Collaboratively Use Hypermedia to Learn Science? The Dynamics of Self-and other-Regulatory Processes in an Ecology Classroom: *Http://Dx.Doi.Org/10.2190/HFT6-8EB1-TN99-MJVQ*, *31*(3), 215–245. https://doi.org/10.2190/HFT6-8EB1-TN99-MJVQ

Bean, J. P., & Metzner, B. S. (1985). A Conceptual Model of Nontraditional Undergraduate Student Attrition Author (s): John P . Bean and Barbara S . Metzner Published by : American Educational Research Association Stable URL : https://www.jstor.org/stable/1170245 REFERENCES Linked reference. *American Educational Research Journal*, *55*(4), 485–540.

Berger, R. (2015). Now I see it, now I don't: researcher's position and reflexivity in qualitative research. *Qualitative Research*, *15*(2), 219–234. https://doi.org/10.1177/1468794112468475

Billings, C. E., Lauber, J. K., Funkhouser, H., Lyman, E. G., & Huff, E. M. (1976). *NASA aviation safety reporting system quarterly report*. *76*, 21.

Bing, L., Yaoguang, Q., & Jiyun, D. (2014). Time-history simulation of civil architecture earthquake disaster relief- based on the three-dimensional dynamic finite element method. *Frattura Ed Integrita Strutturale*, *30*, 526–536. https://doi.org/10.3221/IGF-ESIS.30.63

Boone, H. N., & Boone, D. A. (2012). Analyzing likert data. *Journal of Extension*, *50*(2), 1–5. https://www.researchgate.net/profile/Mahdi-Safarpour-2/post/what_is_a_logistic_regression_analysis/attachment/59d622fb79197b8077981513/AS%3A304626539139073%401449640034657/download/Likert+Scale+vs+Likert+Item+%28Good+Source%29.pdf

Boynton, M. H., Portnoy, D. B., & Johnson, B. T. (2015). Exploring the ethics and psychological impact of deception in psychological research. *IRB*, *35*(2), 7–13. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4502434/

Brophy, S. P., Magana, A. J., & Strachan, A. (2013). Lectures and simulation laboratories to improve learners' conceptual understanding. *Advances in Engineering Education*, *3*(3), 1–28.

Brown, M. (Dartmouth C. (2009). Learning and technology, in that Order. *EDUCAUSE Review*, *44*(4 (July/August 2009)), 62–63.

Bryant, S. L. J. (2011). *The dangers of an over-reliance on technology*.

Campbell, E. M., Sittig, D. F., Guappone, K. P., Dykstra, R. H., & Ash, J. S. (2007). Overdependence on technology: an unintended adverse consequence of computerized provider order entry. *AMIA ... Annual Symposium Proceedings / AMIA Symposium. AMIA Symposium*, 94–98.

Chaturvedi, S., Prabhakaran, R., Yoon, J., & Abdel-Salam, T. (2011). Advances in Engineering Education Engineering Laboratory Instruction in Virtual Department of Mechanical and Aerospace Engineering. *Advances in Engineering Education*, *2*(4), 1–24.

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A Partial Least Squares Latent Variable Modeling Approach for Measuring Interaction Effects: Results from a Monte Carlo Simulation Study and an Electronic-Mail Emotion/Adoption Study. In *Research* (Vol. 14, Issue 2). https://www.jstor.org/stable/23011467

Chyung, S. Y. (Yonnie), Moll, A., Marx, B., Frary, M., & Callahan, J. (2010). Improving engineering students' cognitive and affective preparedness with a pre-instructional E-learning strategy. *Advances in Engineering Education*, *2*(1), 1–28.

Clubb, G. (2010). The Sensor Irony : How Reliance on Sensor Technology. *Work*, *0704*.

Cohen, B. H. (2008). *Explaining psychological statistics*. John Wiley & Sons.

Coulehan, M. B., & Wells, J. F. (2012). Guidelines for responsible data management in scientific research. *(U.S. Department of Health and Human Services Office of Research Integrity)*, 1–46.

Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.

Cronbach, L. J. (1951). COEFFICIENT ALPHA AND THE INTERNAL STRUCTURE OF TESTS*. In *PSYCHOMETRIKA* (Vol. 16, Issue 3).

Dabbagh, N., & Beattie, M. (2010). Student and instructor perceptions of the usefulness of computer-based microworlds in supporting the teaching and assessment of computer networking skills: An exploratory study. *Advances in Engineering Education*, *2*(1), 1–27.

Dalcher, D. (2007). Why the pilot cannot be blamed: A cautionary note about excessive reliance on technology. *International Journal of Risk Assessment and Management*, *7*(3), 350–366. https://doi.org/10.1504/IJRAM.2007.011988

Dang, W., Konietzky, H., Herbst, M., & Frühwirt, T. (2017). Complex analysis of shear box tests with explicit consideration of interaction between test device and sample. *Measurement: Journal of the International Measurement Confederation*, *102*, 1–9. https://doi.org/10.1016/j.measurement.2017.01.040

Davidovitch, L., Parush, A., & Shtub, A. (2006). Simulation-based Learning in Engineering Education: Performance and Transfer in Learning Project Management. *Journal of Engineering Education*, *95*(4), 289–299. https://doi.org/10.1002/j.2168-9830.2006.tb00904.x

Davis, D., & Follette, W. (2001). Foibles of Witness Memory for Traumatic/High Profile Events. *Journal of Air Law and Commerce*, *66*(4), 1421.

Deley, T., & Dubois, E. (2020). Assessing Trust Versus Reliance for Technology Platforms by Systematic Literature Review. *Social Media and Society*, *6*(2). https://doi.org/10.1177/2056305120913883

Dukes, R. L., & Waller, S. J. (1976). Toward a general evaluation model for simulation games: Gem. In *Simulation & Gaming* (Vol. 7, Issue 1, pp. 75–88). https://doi.org/10.1177/104687817600700106

Erdfelder, E., Faul, F., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fedorova, N. N., Valger, S. A., & Fedorov, A. v. (2016). Simulation of blast action on civil structures using ANSYS Autodyn. *AIP Conference Proceedings*, *1770*(October 2016). https://doi.org/10.1063/1.4963939

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. In *Source: Journal of Marketing Research* (Vol. 18, Issue 1).

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational Research: an introduction*. Pearson.

Gero, A., Zoabi, W., & Sabag, N. (2014). Animation based learning of electronic devices. *Advances in Engineering Education*, *4*(1), 1–22.

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, *19*(1), 121–127. https://doi.org/10.1136/amiajnl-2011-000089

Goddard, K., Roudsari, A., & Wyatt, J. C. (2014). Automation bias: Empirical results assessing influencing factors. *International Journal of Medical Informatics*, *83*(5), 368–375. https://doi.org/10.1016/j.ijmedinf.2014.01.001

Goeser, P. T., Johnson, W. M., Hamza-Lup, F. G., & Schaefer, D. (2011). VIEW - A virtual interactive web-based learning environment for engineering. *Advances in Engineering Education*, *2*(3), 1–25.

Grant, A., Williams, P., Ward, N., & Basker, S. (2009). GPS jamming and the impact on maritime navigation. *Journal of Navigation*, *62*(2), 173–187. https://doi.org/10.1017/S0373463308005213

Grissinger, M. (2019). Understanding human over-reliance on technology. *P and T*, *44*(6), 320–321.

Hohenwarter, M., Borcherds, M., Ancsin, G., Bencze, B., Blossier, M., Elias, J., Frank, K., Gal, L., Hofstaetter, A., Jordan, F., Karacsony, B., Konecny, Z., Kovacs, Z., Kuellinger, W., Lettner, E., Lizelfelner, S., Parisse, B., Solyom-Gecse, C., & Tomaschko, M. (2018). *GeoGebra 6.0.507.0*.

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using Mixed-Methods Sequential Explanatory Design: From Theory to Practice. *Field Methods*, *18*(1), 3–20. https://doi.org/10.1177/1525822X05282260

Jacob, S. A., & Furgerson, S. P. (2012). Writing Interview Protocols and Conducting Interviews: Tips for Students New to the Field of Qualitative Research. *The Qualitative Report*, *17*(T & L Art. 6), 1–10. https://doi.org/10.1016/0168-1702(91)90033-r

Javadi, A. H., Emo, B., Howard, L. R., Zisch, F. E., Yu, Y., Knight, R., Pinelo Silva, J., & Spiers, H. J. (2017). Hippocampal and prefrontal processing of network topology to simulate the future. *Nature Communications*, *8*. https://doi.org/10.1038/ncomms14652

Kauffman, D. F. (2016). Self-Regulated Learning in Web-Based Environments: Instructional Tools Designed to Facilitate Cognitive Strategy Use, Metacognitive Processing, and Motivational Beliefs: *Http://Dx.Doi.Org/10.2190/AX2D-Y9VM-V7PX-0TAD*, *30*(1–2), 139–161. https://doi.org/10.2190/AX2D-Y9VM-V7PX-0TAD

Kemmelmeier, M., Davis, D., & Follette, W. C. (2003). Seven "sins" of misdirection? Ethical controversies surrounding the use of deception in research. In *Handbook of Professional Ethics for Psychologists: Issues, Questions, and Controversies* (Issue January). https://doi.org/10.4135/9781412990004.n14

Khan, F., & Singh, K. (2015). Curricular improvements through computation and experiment based learning modules. *Advances in Engineering Education*, *4*(4), 1–20.

Koh, C., Tan, H. S., Tan, K. C., Fang, L., Fong, F. M., Kan, O., Lin Lye, S., & Lin Wee, A. (2010). Investigating the effect of 3D simulation-based learning on the motivation and performance of engineering students. *Journal of Engineering Education*, *99*(3), 237–251. https://doi.org/10.1002/j.2168-9830.2010.tb01059.x

Kolari, S., & Savander-Ranne, C. (2000). Will the Application of Constructivism Bring a Solution to Today's Problems of Engineering Education? *Global Journal of Engineering Education*, *4*(3), 275–280.

Kollöffel, B., & de Jong, T. A. J. M. (2013). Conceptual understanding of electrical circuits in secondary vocational engineering education: Combining traditional instruction with inquiry learning in a virtual lab. *Journal of Engineering Education*, *102*(3), 375–393. https://doi.org/10.1002/jee.20022

Kos, S., Brčić, D., & Musulin, I. (2013). Smartphone application GPS performance during various space weather conditions: A preliminary study. *Conference: International Symposium on Electronics in Transport (ISEP)*, *21*(May 2014). https://www.researchgate.net/publication/236684692_Smartphone_application_GPS_performance_during_various_space_weather_conditions_A_preliminary_study

Li, F., Peng, W., Huang, C. T., & Zou, X. (2013). Smartphone strategic sampling in defending enterprise network security. *IEEE International Conference on Communications*, 2155–2159. https://doi.org/10.1109/ICC.2013.6654846

Lindsay, E. D., & Wankat, P. C. (2012). Going the way of the slide rule: Can remote laboratories fungibly replace the in-person experience? *International Journal of Engineering Education*, *28*(1), 192–201.

Lou, Y., & Sun, Q. (2021). Over-reliance on database: A case study of using web of science. *Human Behavior and Emerging Technologies*, *May*, 1–6. https://doi.org/10.1002/hbe2.266

Magana, A. J. (2017). Modeling and Simulation in Engineering Education: A Learning Progression. *Journal of Professional Issues in Engineering Education and Practice*, *143*(4). https://doi.org/10.1061/(ASCE)EI.1943-5541.0000338

Magana, A. J., Brophy, S. P., & Bodner, A. M. (2012). Instructors' intended learning outcomes for using computational simulations as learning tools. *Journal of Engineering Education*, *101*(2), 220–243. https://doi.org/10.1002/j.2168-9830.2012.tb00049.x

McCullough, D., & Collins, R. (2019). “Are we losing our way?” Navigational aids, socio-sensory way-finding and the spatial awareness of young adults. *Area*, *51*(3), 479–488. https://doi.org/10.1111/area.12478

McHenry, A. L., Depew, D. R., Dyrenfurth, M. J., Dunlap, D. D., Keating, D. A., Stanford, T. G., Lee, P., & Deloatch, G. (2005). Constructivism: The learning theory that supports competency development of engineers for engineering practice and technology leadership through graduate education. *ASEE Annual Conference and Exposition, Conference Proceedings*, *January 2015*, 2263–2268.

Meschke, G., Neu, G. E., & Marwan, A. (2019). Robust segmental lining design – Potentials of advanced numerical simulations for the design of TBM driven tunnels. *Geomechanik Und Tunnelbau*, *12*(5), 484–490. https://doi.org/10.1002/geot.201900032

Milgram, S. (1963). Behavioral Study of obedience. *Journal of Abnormal and Social Psychology*, *67*(4), 371–378. https://doi.org/10.1037/h0040525

Nahvi, M. (1996). Dynamics of student-computer interaction in a simulation environment: Reflection on curricular issues. *Proceedings - Frontiers in Education Conference*, *3*, 1383–1386. https://doi.org/10.1109/fie.1996.568522

Narciss, S., Proske, A., & Koerndle, H. (2007). Promoting self-regulated learning in web-based learning environments. *Computers in Human Behavior*, *23*(3), 1126–1144. https://doi.org/10.1016/J.CHB.2006.10.006

National Science Foundation. (2021). *Research on Emerging Technologies for Teaching and Learning PROGRAM SOLICITATION NSF 20-612 REPLACES DOCUMENT (S):* 1–11.

Navaee, S., & Kang, J. (2017). Finite element simulation models for mechanics of materials. *ASEE Annual Conference and Exposition, Conference Proceedings*, *2017-June*. https://doi.org/10.18260/1-2--28359

Nikolic, D., Jaruhar, S., & Messner, J. I. (2011). Educational Simulation in Construction: Virtual Construction Simulator. *Journal of Computing in Civil Engineering*, *25*(6), 421–429. https://doi.org/10.1061/(asce)cp.1943-5487.0000098

Olivas, L., & Newstrom, J. W. (1981). Learning through the use of simulation games. In *Training and Development Journal* (pp. 63–66).

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. In *Frontiers in Psychology* (Vol. 8, Issue APR). Frontiers Media S.A. https://doi.org/10.3389/fpsyg.2017.00422

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, *52*(3), 381–410. https://doi.org/10.1177/0018720810376055

Public Health. (n.d.). *Why Mixed Methods*. Johns Hopkins Bloomberg School of Public Health. Retrieved October 31, 2021, from https://publichealth.jhu.edu/academics/program-finder/training-grants/mixed-methods-research-training-program-for-the-health-sciences/about-the-program/why-mixed-methods#cite-02

Riley, D. M. (2014). What's wrong with evidence? Epistemological roots and pedagogical implications of "evidence-based practice" in STEM education. *ASEE Annual Conference and Exposition, Conference Proceedings*. https://doi.org/10.18260/1-2--23306

Rivers, D. J., Nakamura, M., & Vallance, M. (2022). Online Self-Regulated Learning and Achievement in the Era of Change. *Journal of Educational Computing Research*, *60*(1), 104–131. https://doi.org/10.1177/07356331211025108

Rojko, A., Hercog, D., & Jezernik, K. (2010). Mechatronics E-course for regular students and adults: Realization and comparison of efficiency. *2010 IEEE Education Engineering Conference, EDUCON 2010*, 959–966. https://doi.org/10.1109/EDUCON.2010.5492475

Rokooei, S., Goedert, J. D., & Najjar, L. (2017). Enhancing Construction Project Management Education by Simulation. *Journal of Professional Issues in Engineering Education and Practice*, *143*(4), 1–9. https://doi.org/10.1061/(ASCE)EI.1943-5541.0000329

Sauvé, L., Renaud, L., Kaufman, D., & Marquis, J. S. (2007). Distinguishing between games and simulations: A systematic review. *Educational Technology and Society*, *10*(3), 247–256.

Shao, X. F., Ye, L. W., Cai, M. J., & Wang, Y. (2008). Simultaneous image denoising and curve extraction by tensor voting. *Proceedings - International Symposium on Computer Science and Computational Technology, ISCSCT 2008*, *1*, 536–538. https://doi.org/10.1109/ISCSCT.2008.184

Sommers, R., & Miller, F. G. (2013). Forgoing Debriefing in Deceptive Research: Is It Ever Ethical? *Ethics and Behavior*, *23*(2), 98–116. https://doi.org/10.1080/10508422.2012.732505

Sorensen, A. A., & Snider, J. C. (2001). Is increasing our reliance on technology enhancing education? *Journal of Continuing Higher Education*, *49*(1), 11–18. https://doi.org/10.1080/07377366.2001.10400419

Tang, S. (2014). An interactive simulator-based pedagogical (ISP) approach for teaching microcontrollers in engineering programs. *Advances in Engineering Education*, *4*(2), 1–19.

Trevas, D., & Nimkoff, T. (2015). *Mixed Methods in Education Research IES Technical Working Group Meeting Summary*.

Uribe, M. D. R., Magana, A. J., Bahk, J. H., & Shakouri, A. (2016). Computational simulations as virtual laboratories for online engineering education: A case study in the field of thermoelectricity. *Computer Applications in Engineering Education*, *24*(3), 428–442. https://doi.org/10.1002/cae.21721

Uz, I., & Kemmelmeier, M. (2017). Can deception be desirable? *Social Science Information*, *56*(1), 98–106. https://doi.org/10.1177/0539018416675070

van der Meij, J., & de Jong, T. (2006). Supporting students' learning with multiple representations in a dynamic simulation-based learning environment. *Learning and Instruction*, *16*(3), 199–212. https://doi.org/10.1016/j.learninstruc.2006.03.007

Vlachopoulos, D., & Makri, A. (2017). The effect of games and simulations on higher education: a systematic literature review. In *International Journal of Educational Technology in Higher Education* (Vol. 14, Issue 1). International Journal of Educational Technology in Higher Education. https://doi.org/10.1186/s41239-017-0062-1

Wandtner, B. (2018). *Non-driving related tasks in highly automated driving-Effects of task characteristics and drivers' self-regulation on take-over performance Acknowledgement*. Wurzburg.

Wankat, P. C. (2002). Integrating the use of commercial simulators into lecture courses. *Journal of Engineering Education*, *91*(1), 19–23. https://doi.org/10.1002/j.2168-9830.2002.tb00668.x

Wankat, P. C., Felder, R. M., Smith, K. A., & Oreovicz, F. S. (2002). The Scholarship of Teaching and Learning in Engineering. In M. T. Huber & S. Morreale (Eds.), *Disciplinary* *styles in sthe scholarship of teaching and learning: Exploring common ground.* AAHE/Carnegie Foundation for the Advancement of Teaching. https://doi.org/10.1093/jaarel/lfw006

Wankat, P. C., & Oreovics, F. S. (2015). Learning Theories. In *Teaching Engineering* (pp. 357–390).

Xiao, Y., Zhang, J., Cao, J., & Li, C. (2021). Prefabricated Urban Underground Utility Tunnels: A Case Study on Mechanical Behaviour with Strain Monitoring and Numerical Simulation. *Advances in Materials Science and Engineering*, *2021*, 1–14. https://doi.org/10.1155/2021/5534526

Zimmerman, B. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology*, *11*(4), 307–313. https://doi.org/10.1016/0361-476X(86)90027-5

Zimmerman, B., & Moylan, A. (2009). Self-regulation: Where metacognition and motivation intersect. In *Handbook of metacognition in education* (pp. 311–328). Routledge. https://books.google.com/books?hl=en&lr=&id=JpWOAgAAQBAJ&oi=fnd&pg=PA299&dq=zimmerman+2009+self+regulation&ots=eytb3dikzd&sig=pZAsw_07LGe9u70rXsaogqsd8yc#v=onepage&q=zimmerman%202009%20self%20regulation&f=false

APPENDICES

Appendix A

Demographic Questionnaire

[Informed Consent Information will come first – see other document. Quiz will exit if they do not wish to submit their informed consent at this time without gathering any other information.]

Please enter your A number. This is only used to ensure that your data is collected and placed together for proper analysis. It will be converted to an identification number prior to any analysis. (Open Ended)

To which gender identity do you most identify: Male, Female, Non-binary, Prefer not to answer, Prefer to Self Describe (Open Ended)

Are you of Hispanic, Latino, or of Spanish Origin? Yes/No

How would you describe yourself? Hispanic, Black or African American, Asian, Native Hawaiian or Other Pacific Islander, White/Caucasian, Native American, Mixed Racial, International, Prefer not to answer

What is your age group? 17 or less, 18-21, 22-24, 25-27, 28-30, 31-33, 33+

Please select your cumulative GPA: <2, 2.33, 2.66, 3, 3.33, 3.66, 4, other

What is your class? Freshman, Sophomore, Junior, Senior

Which Engineering Program are you currently in or do you plan to pursue? Biological, Civil, Environmental, Electrical, Mechanical, Other

What is your comfort level or familiarity with Educational Software (software used to help you learn a particular topic)? (Likert, 7 point scale)

Hours spent with Educational Software: 0-40, 40-80, 80-120, 120+

Which software have you used in the past? (Open Ended)

Appendix B

Reliance on Technology Questionnaire

[This questionnaire is given as a Qualtrics questionnaire. It was adapted from Al-Natour et al., (2008) for the Likert questions. The internal validity and reliability are discussed in Chapter IV under the Internal Reliability of Questionnaire section of this dissertation. The first five areas are Likert scale questions. The last few areas are open ended questions related to the actual nature of the study. Information found inside [] is intended to explain the purpose of the section and will not be in the Qualtrics questionnaire.]

Thank you for participating in this questionnaire. The first question will gather your A Number so that your results can be combined with the first questionnaire and with your simulation results. After your data is gathered together, your A Number will be removed from the data to protect your anonymity. The next five questions are scaled questions that will address your use of the simulation. Answer them to the best of your ability.

1. What is your A Number?

Reuse Intention	**Degree to I agree with the statement**						
	Strongly Disagree	Moderately Disagree	Slightly Disagree	Neutral	Slightly Agree	Moderately Agree	Strongly Agree
I intend to reuse the truss solving assistant for the same solving task in the future.	1	2	3	4	5	6	7
I predict that I will reuse the truss solving assistant for the same truss solving task in the future.	1	2	3	4	5	6	7
I would consider using the truss solving assistant for similar future problems.	1	2	3	4	5	6	7
I am willing to use this truss solving assistant as an aid to help solve future problems.	1	2	3	4	5	6	7
I am willing to let this truss solving assistant assist me in solving trusses in the future.	1	2	3	4	5	6	7

Perceived Usefulness	**Degree to I agree with the statement**						
	Strongly Disagree	Moderately Disagree	Slightly Disagree	Neutral	Slightly Agree	Moderately Agree	Strongly Agree
Using the truss solving assistant enabled me to solve the problem more quickly.	1	2	3	4	5	6	7
In my opinion, using the truss solving assistant increased my solving effectiveness.	1	2	3	4	5	6	7
In my opinion, using the truss solving assistant increased my solving efficiency.	1	2	3	4	5	6	7
Overall, the truss solving assistant was useful for solving truss problems.	1	2	3	4	5	6	7

Degree to I agree with the statement

Trust	Strongly Disagree	Moderately Disagree	Slightly Disagree	Neutral	Slightly Agree	Moderately Agree	Strongly Agree
I believe this truss solving assistant is competent.	1	2	3	4	5	6	7
I believe this truss solving assistant is benevolent.	1	2	3	4	5	6	7
I believe this truss solving assistant has high integrity.	1	2	3	4	5	6	7
Overall, I believe this truss solving assistant is trustworthy.	1	2	3	4	5	6	7

Perceived Process Similarity:	**Degree to I agree with the statement**						
How similar or different do you think you and the truss solving assistant are in terms of:	Very Different	Mostly Different	More Different than the Same	Somewhat the Same	More the Same than Different	Mostly the Same	Exactly the Same
Your problem-solving style.	1	2	3	4	5	6	7
The order in which you solve problems.	1	2	3	4	5	6	7
How you arrived at a final answer for the truss.	1	2	3	4	5	6	7

Domain Knowledge	**Degree to I agree with the statement**						
	Strongly Disagree	Moderately Disagree	Slightly Disagree	Neutral	Slightly Agree	Moderately Agree	Strongly Agree
I consider myself to be an expert in solving truss problems.	1	2	3	4	5	6	7
I am knowledgeable about trusses.	1	2	3	4	5	6	7
I have extensive experience in solving truss problems.	1	2	3	4	5	6	7

[Information found in {} are updated based on relevant information based on which group and which error type, if appropriate, the student encountered. Students will only see one version here. A Priori codes that the questions are shown in parenthesis following each question. The participants will not see these codes while taking the questionnaire.]

Initial Simulation Responses

The following questions relate to your experience using the simulation. Please answer them to the best of your ability. If you feel that you cannot answer a particular question, please feel free to leave it blank.

1) How do you feel about using guided simulations to help learn a topic? (Trust/Help)

2) Did you need to recheck your answer because the program said it disagreed with you? If so, how many times did you recheck your work? (Complacency/Vigilance)
3) Did you change your answer when given the opportunity by the simulation? If so, why? (Switch)
4) How did you check your answer before resubmitting it? (Vigilance/Trust)
5) How often did you use the help provided by the simulation? (Help)
6) Did the help provided by the simulation help you solve the truss problem? If so, how? (Help/Trust)
7) Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not? (Complacency/Awareness)
8) What did you do if you disagreed with the simulation? (Vigilance/Switch)
9) What was the biggest impact that the guided simulation had on your understanding of the Method of Joints? (Trust)
10) What would you change about the guided simulation?

Updated Purpose of the Study *(Used to reveal the deception to participants)*

This study was actually measuring the extent to which students rely on technology when learning new material. It was intended to gather information about what your initial and final calculations are for each force in the system, the confidence you had in your initial calculations, the number of times you opted to use the help provided by the system, and how often you switched your answers. More information can be found in the update Conflict of Interest found below.

We believe that this deception was necessary to help us learn more about student reliance on technology, an especially important area given the recent online educational experiences student have had due to Covid. If you had known the true nature of the study, you may not have behaved normally in interacting with the simulation, which would make it harder to measure this reliance. We apologize if you feel that any harm has happened through this process. Your participation has helped us develop an understanding towards how students may interact with simulations which can be especially beneficial during a forced online teaching experience where students are asked to engage with quickly assembled teaching mediums. Please note that, as before, all information gathered from this study will be held as anonymous as possible. The only reason we are gathering your A number at each step is to make sure we can correlate your three separate data sources (initial, this questionnaire, and your simulation results). Once we have connected your data, we will apply a random identifier to your data to keep the analysis anonymized – you're A number will be removed from the questionnaire results. We encourage you to continue participating in this study, but want to recognize that you get to decide if you are ok participating in this study at all. Please note that if you choose to withdraw from the study, we will not use any of your results from any questionnaire or from the simulation in our analysis. If you have questions or concerns about the deceptive nature of this study, please contact Jon Anderson at [enter USU email here].

To help in gathering the information that would help us research student reliance on technology, you were placed into one of four groups, including three treatment groups and a control group. The first group, the control group, was provided with a correctly coded simulation and were not informed that there may be potential errors. The second group was provided with a faulty simulation, which had an incorrect process coded into it, and were not informed that there were any potential errors available. The third group was provided with a correctly coded simulation but were informed that there may be errors in the simulation. The fourth group was provided with an incorrectly coded simulation, like the first treatment group, but were informed that there might be errors found in the simulation. The code you were provided with told the simulation which of the following four solution patterns to apply to your experience. Groups one and three were provided with one of the first two solution patterns, while groups two and four were provided one of the last two solution patterns.

1. The correct methods for finding the solution
2. Changing the signs of Tension/Compression such that Tension was negative and Compression was positive
3. Swapping the SIN/COS calculations from Vertical and Horizontal Force Analysis – the fraction used to calculate values was inverted
4. Doing both 2 and 3 at the same time

In all cases, consistent answers were reached. That is, the simulation had a consistent answer that worked for all members. This allowed you to work through the truss and, if following the advice of the program, your last step should show that the forces were in balance. However, students with patterns 3 and 4 would lead to incorrect forces at each member, even if they were consistent as a whole.

In the process of working with the simulation, you were assigned to {list appropriate treatment group and solution pattern}, which impacted the help and the answer check at the end of each step. As such, the system recorded what each answer was, whether you changed your answer, and your confidence level in each initial answer and in the process as a whole.

This data is being collected and analyzed to determine what factors play a role in students over-relying on technology. Your answers to the following questions will help outline some of these factors and how you feel about them.

[The updated Informed Consent document will be included here, including contact information for the researcher (Jon Anderson) if the participant has further questions.]

1) Now that you know the true purpose of the study, do you wish to continue your participation in this study at this time? Note that if you answer yes, your data collected from all aspects of this study will be included in the study, but if your answer is no, your data will not be included. {If no, the questionnaire will end here and all data gathered by this student will be excluded from the study}
2) How do you feel about the fact that the purpose of the study was not shared from the beginning? (Trust)
3) Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect? (Trust/Vigilance)
4) How does knowing the information above about the intentional errors impact your thoughts on the simulation? (Trust/Complacency/Vigilance)
5) Has your perception of the help provided by the simulation changed after learning the true nature of the study and the simulation? (Help, Trust)
6) How do you feel about having the simulation try to correct your answers now that you know it may have compared your answer with an incorrect one? (Complacency/Vigilance)
7) What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors? (Vigilance/Trust)
8) How does this study impact your perceived usage of guided simulations moving forward? (Trust)

9) When is the ideal time to know that a guided simulation or other technology might be faulty? (Trust)
10) How would you ensure that a guided simulation or other technology you are using isn't faulty? (Vigilance)
11) Are you interest in potentially participating in an interview to further explore these topics? Only students who answer yes to this question will be considered for the interview but answering yes is not a guarantee that you will be selected.

Appendix C

Interview Protocol

This interview protocol is intended to delve deeper into responses found in the Reliance on Technology Questionnaire. The goal was to find two participants from each treatment group to be interviewed. This protocol provides the overview of the interview. The interview is intended to be completed in about half an hour, give or take the time the participant is spending answering each question. Items in quotation marks are intended to be a source of direct quotation, although the particular words for any single interview may be changed. They serve as the guide for the interview, shaping the direction of the conversation. Bullet point items are meant to serve as prompts for the interviewer, making sure that key points of the research get asked. Also, material found in [] is intended to show what A Priori codes the topic is intended to find out about.

The

Script:

"Hello, my name is Jon Anderson. I'm a Doctoral Candidate in the Department of Engineering Education, and I have been an associate professor of mathematics at Utah Valley University since 2012. I have always enjoyed using appropriate technology in my classes, and that's part of why I wanted to explore the uses and reliance on these technologies in my dissertation."

"As a reminder, you have earlier provided consent to participate in this interview by filling out the Informed Consent form as part of the original questionnaire. You are welcome to withdraw your consent at any time. If this happens during the interview, please inform me that you would like to stop, and we will end the interview and I will remove the recordings. You do not have to remove all of your consent at this time – you have control over whether you would like to stop this interview and remove it in its entirety or in part."

"This interview is being recorded, but the recording will only be available to me, the researcher. Once the video has been transcribed, any references to you that could be used to identify you will be anonymized, and the transcription will be kept in a different location. Direct quotes from your interview may be used, but will be attributed to the anonymized participant. Most of the interview, however, will be analyzed looking for themes that can be written about more generally."

"If you wish to continue, please tell me about yourself, including your use of technology in your learning and with solving trusses using the Method of Joints." [Complacency]

Use of technological learning aids
Familiarity with truss problems

"Tell me about your experience using the guided simulation." [Trust, Complacency, Vigilance]

Instructions in use
Use of help tools
Confidence in work
Was there a "personal" connection with the simulation (flow of words, first person responses, etc.)

"What was your strategy in solving the provided truss problem?" [Trust]

Verification of work (rely on simulation, do manually, etc.)
Order of joints visited

"How often did the simulation disagree with your answer for a force when you first entered it?" [Complacency, Switch, Vigilance]

"When the simulation disagreed with your answer, what did you do?" [Switch, Trust]

Feelings about the disagreement
How they handled the disagreement
How often did they question the simulation

"What was your experience using the help provided by the simulation?" [Help]

Trust given to simulation help
Level of help sought for
Helpfulness of the help

"How did you feel when you found out the guided simulation might intentionally contain errors?" [Complacency, Trust]

When did the participant find out
Initial responses
Thoughts after having some time away from it
Affect the help/verification process?

"Other than removing the intended errors in the program, how can the guided simulation be improved to help you better understand the method of joints analysis of a truss?"

Visual/written help
Visual representations
Feedback

"Is there anything else you would like to mention in relation to the guided simulation or the study as a whole?"

"Thank you for participating in the study. Once I have transcribed your interview, I will perform a basic analysis of it and attempt to highlight key findings from it. I may also need to reach out to clarify something from the interview. I will also only talk about this interview with you – I will not be sharing portions of this interview with others or their interviews with you. Can I share my analysis of your interview with you to either clarify something you said and make sure that I accurately portrayed your thoughts?"

"This will happen through email from, and you will have the opportunity to respond. If you have further questions about the study or your part in it, please reach out to me at the same email."

Appendix D

Tables relating to Automation Complacency Qualitative Coding of Questionnaire and Interviews

This appendix contains the tables related to each of the Reliance on Technology questionnaire and the interviews, separating out the grouping categories for individual reporting. Table 45 shows the frequency with which certain codes showed up in participant responses to the questions from the questionnaires and interviews divided into whether the participant had a correctly or an incorrectly working simulation. Table 46 shows the frequency with which certain codes showed up in participant responses to the questions from the questionnaires and interviews divided into which simulation group the participant belongs to. Table 47 shows the frequency with which certain codes showed up in participant responses to the questions from the questionnaires and interviews divided by participant confidence. Note that the interviews were excluded from Table 47 because all four interviewees were in the high confidence group. Table 48 shows the frequency with which certain codes showed up in participant responses to the questions from the questionnaires and interviews divided by participant experience. Table 49 shows the frequency with which certain codes showed up in participant responses to the questions from the questionnaires and interviews divided by participant experience. Note that the interviews were excluded from Table 49 because all four interviewees were in the low experience group.

Qualitative Coding of Automation Complacency based on Correctness of Simulation

Table 45

Qualitative Coding of Automation Complacency Questions, Correctness of Simulation

Question	Coding	Correctness	Count	%
Did you need to recheck your answer because the program said it disagreed with you?	Complacency	Correct	15	88.2
		Incorrect	15	88.2
	Vigilance	Correct	0	0
		Incorrect	2	11.8
Did you change your answer when given the opportunity to do so by the simulation? If so, why?	Switch	Correct	14	82.4
		Incorrect	15	88.2
	Complacency[a]	Correct	1	5.9
		Incorrect	0	0
	Vigilance[a]	Correct	0	0
		Incorrect	4	23.5
How did you check your answer before submitting it?	Vigilance	Correct	0	0
		Incorrect	1	5.9
	Complacency[a]	Correct	15	88.2
		Incorrect	13	76.4
	Switch	Correct	1	5.9
		Incorrect	2	11.8
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Complacency	Correct	14	82.4
		Incorrect	13	76.5
What did you do if you disagreed with the simulation?	Vigilance	Correct	2	11.8
		Incorrect	7	41.2
	Switch	Correct	3	17.6
		Incorrect	7	41.2
	Complacency[a]	Correct	6	35.3
		Incorrect	6	35.3
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Vigilance	Correct	8	47.1
		Incorrect	13	76.5
How does knowing about the intentional errors impact your thoughts on the simulation?	Complacency	Correct	4	23.5
		Incorrect	7	41.2
	Vigilance	Correct	5	29.4
		Incorrect	5	29.4
	Switch	Correct	1	5.9
		Incorrect	0	0

Question	Coding	Correctness	Count	%
How do you feel about having the simulation try to correct your answers now that you know it may have compared your answers with an incorrect one?	Complacency	Correct	12	70.6
		Incorrect	11	64.7
	Vigilance	Correct	0	0
		Incorrect	4	23.5
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Vigilance	Correct	1	5.9
		Incorrect	0	0
	Switch	Correct	0	0
		Incorrect	1	5.9
How would you ensure that a guided simulation or other technology you are using isn't faulty?	Vigilance	Correct	10	58.8
		Incorrect	9	52.9
Tell me about yourself, including your use of technology in your learning and with solving trusses using the Method of Joints.	Complacency	Correct	1	100
		Incorrect	3	100
Tell me about your experience using the guided simulation.	Complacency	Correct	0	0
		Incorrect	1	33.3
	Vigilance	Correct	0	0
		Incorrect	1	33.3
How often did the simulation disagree with your answer for a force when you first entered it?	Complacency	Correct	1	100
		Incorrect	2	66.7
	Vigilance	Correct	0	0
		Incorrect	2	66.7
	Switch[a]	Correct	1	100
		Incorrect	0	0
When the simulation disagreed with your answer, what did you do?	Switch	Correct	0	0
		Incorrect	2	66.7
	Complacency	Correct	1	100
		Incorrect	1	33.33
	Vigilance	Correct	0	0
		Incorrect	2	66.7
How did you feel when you found out the guided simulation might intentionally contain errors?	Complacency	Correct	1	100
		Incorrect	1	33.3
	Vigilance[a]	Correct	0	0
		Incorrect	1	33.3

[a]Code was an emergent code added while analyzing the data.

Qualitative Coding of Automation Complacency based on Simulation Group

Table 46

Qualitative Coding of Automation Complacency Questions, Simulation Group

Question	Coding	Broad Group	Count	%
Did you need to recheck your answer because the program said it disagreed with you?	Complacency	Control	9	100
		Deceived	7	77.8
		Informed	6	75.0
		Misled	8	100
	Vigilance	Control	0	0
		Deceived	1	11.1
		Informed	0	0
		Misled	1	11.1
Did you change your answer when given the opportunity to do so by the simulation? If so, why?	Switch	Control	8	88.9
		Deceived	9	100
		Informed	6	75.0
		Misled	6	75.0
	Complacency[a]	Control	0	0
		Deceived	0	0
		Informed	1	11.1
		Misled	0	0
	Vigilance[a]	Control	0	0
		Deceived	2	22.2
		Informed	0	0
		Misled	2	25.0
How did you check your answer before submitting it?	Vigilance	Control	0	0
		Deceived	0	0
		Informed	0	0
		Misled	1	12.5
	Complacency[a]	Control	9	100
		Deceived	8	88.9
		Informed	6	75.0
		Misled	5	62.5
	Switch	Control	0	0
		Deceived	1	11.1
		Informed	1	12.5
		Misled	1	12.5
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Complacency	Control	7	77.8
		Deceived	7	77.8
		Informed	7	87.5
		Misled	6	75.0

Question	Coding	Broad Group	Count	%
What did you do if you disagreed with the simulation?	Vigilance	Control	1	11.1
		Deceived	2	22.2
		Informed	1	12.5
		Misled	5	62.5
	Switch	Control	1	11.1
		Deceived	4	44.4
		Informed	1	12.5
		Misled	3	37.5
	Complacency[a]	Control	3	33.3
		Deceived	6	66.7
		Informed	3	37.5
		Misled	0	0
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Vigilance	Control	3	33.3
		Deceived	7	77.8
		Informed	5	62.5
		Misled	6	75.0
How does knowing about the intentional errors impact your thoughts on the simulation?	Complacency	Control	2	22.2
		Deceived	4	44.4
		Informed	2	25.0
		Misled	3	37.5
	Vigilance	Control	1	11.1
		Deceived	3	33.3
		Informed	4	50.0
		Misled	2	25.0
	Switch	Control	1	11.1
		Deceived	0	0
		Informed	0	0
		Misled	0	0
How do you feel about having the simulation try to correct your answers now that you know it may have compared your answers with an incorrect one?	Complacency	Control	8	88.9
		Deceived	6	66.7
		Informed	4	50.0
		Misled	5	62.5
	Vigilance	Control	0	0
		Deceived	2	22.2
		Informed	0	0
		Misled	2	25.0
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Vigilance	Control	1	11.1
		Deceived	0	0
		Informed	0	0
		Misled	1	12.5
	Switch	Control	0	0
		Deceived	1	11.1
		Informed	0	0
		Misled	0	0
	Vigilance	Control	5	55.6%

Question	Coding	Broad Group	Count	%
How would you ensure that a guided simulation or other technology you are using isn't faulty?		Deceived	4	44.4
		Informed	5	62.5
		Misled	5	62.5
Tell me about yourself, including your use of technology in your learning and with solving trusses using the Method of Joints.	Complacency	Control	1	100
		Deceived	1	100
		Informed	0	0
		Misled	2	100
Tell me about your experience using the guided simulation.	Complacency	Control	0	0
		Deceived	0	0
		Informed	0	0
		Misled	1	100
	Vigilance	Control	0	0
		Deceived	0	0
		Informed	0	0
		Misled	0	0
How often did the simulation disagree with your answer for a force when you first entered it?	Complacency	Control	1	100
		Deceived	1	100
		Informed	0	0
		Misled	1	50.0
	Vigilance	Control	0	0
		Deceived	0	0
		Informed	0	0
		Misled	2	100
	Switch[a]	Control	1	100
		Deceived	0	0
		Informed	0	0
		Misled	0	0
When the simulation disagreed with your answer, what did you do?	Switch	Control	0	0
		Deceived	1	100
		Informed	0	0
		Misled	1	50.0
	Complacency	Control	1	100
		Deceived	1	100
		Informed	0	0
		Misled	0	0
	Vigilance	Control	0	0
		Deceived	0	0
		Informed	0	0
		Misled	2	100

Question	Coding	Broad Group	Count	%
How did you feel when you found out the guided simulation might intentionally contain errors?	Complacency	Control	1	100
		Deceived	1	100
		Informed	0	0
		Misled	0	0
	Vigilance[a]	Control	0	0
		Deceived	0	0
		Informed	0	0
		Misled	1	50.0

[a]Code was an emergent code added while analyzing the data.

Qualitative Coding of Automation Complacency based on Confidence

Table 47

Qualitative Coding of Automation Complacency Questions, Confidence of Participant

Question	Coding	Confidence	Count	%
Did you need to recheck your answer because the program said it disagreed with you?	Complacency	High	28	82.4
		Low	2	66.7
	Vigilance	High	2	6.5
		Low	0	0
Did you change your answer when given the opportunity to do so by the simulation? If so, why?	Switch	High	26	83.9
		Low	3	100
	Complacency[a]	High	1	3.2
		Low	0	0
	Vigilance[a]	High	4	12.9
		Low	0	0
How did you check your answer before submitting it?	Vigilance	High	1	3.2
		Low	0	0
	Complacency[a]	High	27	87.1
		Low	1	33.3
	Switch	High	2	6.5
		Low	1	33.3
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Complacency	High	24	77.4
		Low	3	100
What did you do if you disagreed with the simulation?	Vigilance	High	9	29.0
		Low	0	0
	Switch	High	8	25.8
		Low	2	66.7
	Complacency[a]	High	11	35.5
		Low	1	33.3
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Vigilance	High	20	64.5
		Low	1	33.3
How does knowing about the intentional errors impact your thoughts on the simulation?	Complacency	High	10	32.3
		Low	1	33.3
	Vigilance	High	10	32.3
		Low	0	0
	Switch	High	1	3.2
		Low	0	0

Question	Coding	Confidence	Count	%
How do you feel about having the simulation try to correct your answers now that you know it may have compared your answers with an incorrect one?	Complacency	High	21	67.7
		Low	2	66.7
	Vigilance	High	4	12.9
		Low	0	0
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Vigilance	High	2	6.5
		Low	0	0
	Switch	High	1	3.2
		Low	0	0
How would you ensure that a guided simulation or other technology you are using isn't faulty?	Vigilance	High	17	54.8
		Low	2	66.7

[a]Code was an emergent code added while analyzing the data

Qualitative Coding of Automation Complacency based on Trust

Table 48

Qualitative Coding of Automation Complacency Questions, Trust of Participant

Question	Coding	Trust	Count	%
Did you need to recheck your answer because the program said it disagreed with you?	Complacency	High	18	90
		Neutral	2	66.7
		Low	10	90.9
	Vigilance	High	0	0
		Neutral	0	0
		Low	2	18.2
Did you change your answer when given the opportunity to do so by the simulation? If so, why?	Switch	High	17	85.0
		Neutral	3	100
		Low	9	81.8
	Complacency[a]	High	1	5.0
		Neutral	0	0
		Low	0	0
	Vigilance[a]	High	0	0
		Neutral	0	0
		Low	4	36.4
How did you check your answer before submitting it?	Vigilance	High	0	0
		Neutral	0	0
		Low	1	9.1
	Complacency[a]	High	17	85.0
		Neutral	3	100
		Low	8	72.7
	Switch	High	2	10.0
		Neutral	0	0
		Low	1	9.1
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Complacency	High	16	80.0
		Neutral	2	66.7
		Low	9	81.8
What did you do if you disagreed with the simulation?	Vigilance	High	3	15.0
		Neutral	0	0
		Low	6	54.5
	Switch	High	4	20.0
		Neutral	1	33.3
		Low	5	45.5
	Complacency[a]	High	7	35.0
		Neutral	1	33.3
		Low	4	36.4

Question	Coding	Trust	Count	%
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Vigilance	High	10	50.0
		Neutral	2	66.7
		Low	9	81.8
How does knowing about the intentional errors impact your thoughts on the simulation?	Complacency	High	6	30.0
		Neutral	3	100
		Low	2	18.2
	Vigilance	High	4	20.0
		Neutral	1	33.3
		Low	5	45.5
	Switch	High	1	5.0
		Neutral	0	0
		Low	0	0
How do you feel about having the simulation try to correct your answers now that you know it may have compared your answers with an incorrect one?	Complacency	High	14	70.0
		Neutral	3	100
		Low	6	54.5
	Vigilance	High	0	0
		Neutral	2	66.7
		Low	2	18.2
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Vigilance	High	1	5.0
		Neutral	0	0
		Low	1	9.1
	Switch	High	0	0
		Neutral	1	33.3
		Low	0	0
How would you ensure that a guided simulation or other technology you are using isn't faulty?	Vigilance	High	10	50.0
		Neutral	3	100
		Low	6	54.5
Tell me about yourself, including your use of technology in your learning and with solving trusses using the Method of Joints.	Complacency	High	1	100
		Neutral	1	100
		Low	2	100
Tell me about your experience using the guided simulation.	Complacency	High	0	0
		Neutral	0	0
		Low	1	50.0
	Vigilance	High	0	0
		Neutral	0	0
		Low	0	0

Question	Coding	Trust	Count	%
How often did the simulation disagree with your answer for a force when you first entered it?	Complacency	High	1	100
		Neutral	1	100
		Low	1	50.0
	Vigilance	High	0	0
		Neutral	0	0
		Low	2	100
	Switch[a]	High	1	100
		Neutral	0	0
		Low	0	0
When the simulation disagreed with your answer, what did you do?	Switch	High	0	0
		Neutral	1	100
		Low	1	100
	Complacency	High	1	100
		Neutral	1	100
		Low	0	0
	Vigilance	High	0	0
		Neutral	0	0
		Low	2	100
How did you feel when you found out the guided simulation might intentionally contain errors?	Complacency	High	1	100
		Neutral	1	100
		Low	0	0
	Vigilance[a]	High	0	0
		Neutral	0	0
		Low	1	50.0

[a]Code was an emergent code added while analyzing the data.

Qualitative Coding of Automation Complacency based on Experience

Table 49

Qualitative Coding of Automation Complacency Questions, Experience of Participant

Question	Coding	Experience	Count	%
Did you need to recheck your answer because the program said it disagreed with you?	Complacency	Very High	3	75.0
		High	4	66.7
		Low	23	95.8
	Vigilance	Very High	0	0
		High	0	0
		Low	2	8.3
Did you change your answer when given the opportunity to do so by the simulation? If so, why?	Switch	Very High	4	100
		High	4	66.7
		Low	21	87.5
	Complacency[a]	Very High	0	0
		High	0	0
		Low	1	4.2
	Vigilance[a]	Very High	0	0
		High	0	0
		Low	4	16.7
How did you check your answer before submitting it?	Vigilance	Very High	0	0
		High	0	0
		Low	1	4.2
	Complacency[a]	Very High	2	50.0
		High	4	66.7
		Low	19	79.2
	Switch	Very High	1	25.0
		High	0	0
		Low	2	8.3
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Complacency	Very High	4	100
		High	5	83.3
		Low	18	75.0
What did you do if you disagreed with the simulation?	Vigilance	Very High	2	50.0
		High	0	0
		Low	7	29.2
	Switch	Very High	2	50.0
		High	0	0
		Low	8	33.3
	Complacency[a]	Very High	2	50.0
		High	2	33.3
		Low	8	33.3

Question	Coding	Experience	Count	%
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Vigilance	Very High	3	75.0
		High	4	66.7
		Low	14	58.3
How does knowing about the intentional errors impact your thoughts on the simulation?	Complacency	Very High	0	0
		High	2	33.3
		Low	9	37.5
	Vigilance	Very High	0	0
		High	3	50.0
		Low	7	29.2
	Switch	Very High	0	0
		High	0	0
		Low	1	4.2
How do you feel about having the simulation try to correct your answers now that you know it may have compared your answers with an incorrect one?	Complacency	Very High	3	75.0
		High	2	33.3
		Low	18	75.0
	Vigilance	Very High	0	0
		High	1	16.7
		Low	3	12.5
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Vigilance	Very High	1	25.0
		High	0	0
		Low	1	4.2
	Switch	Very High	0	0
		High	0	0
		Low	1	4.2
How would you ensure that a guided simulation or other technology you are using isn't faulty?	Vigilance	Very High	4	100
		High	3	50.0
		Low	12	50.0

[a]Code was an emergent code added while analyzing the data.

Qualitative Coding of Automation Complacency based on Gender

Table 50

Qualitative Coding of Automation Complacency, Gender

Question	Coding	Gender	Count	%
Did you need to recheck your answer because the program said it disagreed with you?	Complacency	Male	22	88.0
		Female	8	88.9
	Vigilance	Male	2	8.0
		Female	0	0
Did you change your answer when given the opportunity to do so by the simulation? If so, why?	Switch	Male	20	80.0
		Female	9	100.0
	Complacency[a]	Male	0	0
		Female	1	11.1
	Vigilance[a]	Male	3	12.0
		Female	1	11.1
How did you check your answer before submitting it?	Vigilance	Male	1	4.0
		Female	0	0
	Complacency[a]	Male	20	80.0
		Female	8	88.9
	Switch	Male	3	12.0
		Female	0	0
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Complacency	Male	20	80.0
		Female	7	77.8
What did you do if you disagreed with the simulation?	Vigilance	Male	7	28.0
		Female	2	22.2
	Switch	Male	7	28.0
		Female	3	33.3
	Complacency[a]	Male	9	36.0
		Female	3	33.3
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Vigilance	Male	15	60.0
		Female	6	66.7
How does knowing about the intentional errors impact your thoughts on the simulation?	Complacency	Male	10	40.0
		Female	1	11.1
	Vigilance	Male	7	28.0
		Female	3	33.3
	Switch	Male	0	0
		Female	1	11.1

Question	Coding	Gender	Count	%
How do you feel about having the simulation try to correct your answers now that you know it may have compared your answers with an incorrect one?	Complacency	Male	20	80.0
		Female	3	33.3
	Vigilance	Male	3	12.0
		Female	1	11.1
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Vigilance	Male	1	4.0
		Female	1	11.1
	Switch	Male	1	4.0
		Female	0	0
How would you ensure that a guided simulation or other technology you are using isn't faulty?	Vigilance	Male	15	60.0
		Female	4	44.4

[a]Code was an emergent code added while analyzing the data

Qualitative Coding of Automation Complacency based on Age

Table 51

Qualitative Coding of Automation Complacency, Age

Question	Coding	Age	Count	%
Did you need to recheck your answer because the program said it disagreed with you?	Complacency	18-21	15	83.3
		22-24	12	92.3
		25+	3	100
	Vigilance	18-21	1	5.6
		22-24	1	7.7
		25+	0	0
Did you change your answer when given the opportunity to do so by the simulation? If so, why?	Switch	18-21	16	88.9
		22-24	10	76.9
		25+	3	100
	Complacency[a]	18-21	0	0
		22-24	0	0
		25+	1	33.3
	Vigilance[a]	18-21	3	16.7
		22-24	1	7.7
		25+	0	0
How did you check your answer before submitting it?	Vigilance	18-21	1	5.6
		22-24	0	0
		25+	0	0
	Complacency[a]	18-21	13	72.2
		22-24	12	92.3
		25+	3	100
	Switch	18-21	2	11.1
		22-24	1	7.7
		25+	0	0
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Complacency	18-21	16	88.9
		22-24	9	69.2
		25+	2	66.7
What did you do if you disagreed with the simulation?	Vigilance	18-21	4	22.2
		22-24	4	30.8
		25+	1	33.3
	Switch	18-21	7	38.9
		22-24	2	15.4
		25+	1	33.3
	Complacency[a]	18-21	7	38.9
		22-24	4	30.8
		25+	0	0

Question	Coding	Age	Count	%
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Vigilance	18-21	13	72.2
		22-24	7	53.8
		25+	1	33.3
How does knowing about the intentional errors impact your thoughts on the simulation?	Complacency	18-21	6	33.3
		22-24	5	38.5
		25+	0	0
	Vigilance	18-21	7	38.9
		22-24	3	23.1
		25+	0	0
	Switch	18-21	0	0
		22-24	1	7.7
		25+	0	0
How do you feel about having the simulation try to correct your answers now that you know it may have compared your answers with an incorrect one?	Complacency	18-21	10	55.6
		22-24	11	84.6
		25+	2	66.7
	Vigilance	18-21	4	22.2
		22-24	0	0
		25+	0	0
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Vigilance	18-21	1	5.6
		22-24	0	0
		25+	1	33.3
	Switch	18-21	1	5.6
		22-24	0	0
		25+	0	0
How would you ensure that a guided simulation or other technology you are using isn't faulty?	Vigilance	18-21	9	50
		22-24	9	69.2
		25+	0	0

[a]Code was an emergent code added while analyzing the data

Qualitative Coding of Automation Complacency based on Major

Table 52

Qualitative Coding of Automation Complacency, Major

Question	Coding	Major	Count	%
Did you need to recheck your answer because the program said it disagreed with you?	Complacency	Mechanical	17	81.0
		Civil	9	100
		Other	4	100
	Vigilance	Mechanical	2	9.5
		Civil	0	0
		Other	0	0
Did you change your answer when given the opportunity to do so by the simulation? If so, why?	Switch	Mechanical	16	76.2
		Civil	9	100
		Other	4	100
	Complacency[a]	Mechanical	1	4.8
		Civil	0	0
		Other	0	0
	Vigilance[a]	Mechanical	3	14.3
		Civil	1	11.1
		Other	0	0
How did you check your answer before submitting it?	Vigilance	Mechanical	1	4.8
		Civil	0	0
		Other	0	0
	Complacency[a]	Mechanical	17	81.0
		Civil	7	77.8
		Other	4	100
	Switch	Mechanical	2	9.5
		Civil	1	11.1
		Other	0	0
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Complacency	Mechanical	17	81.0
		Civil	7	77.8
		Other	3	75.0
What did you do if you disagreed with the simulation?	Vigilance	Mechanical	7	33.3
		Civil	2	22.2
		Other	0	0
	Switch	Mechanical	7	33.3
		Civil	1	11.1
		Other	2	50
	Complacency[a]	Mechanical	8	38.1
		Civil	4	44.4
		Other	0	0

Question	Coding	Major	Count	%
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Vigilance	Mechanical	14	66.7
		Civil	5	55.6
		Other	2	50.0
How does knowing about the intentional errors impact your thoughts on the simulation?	Complacency	Mechanical	6	28.6
		Civil	3	33.3
		Other	2	50
	Vigilance	Mechanical	5	23.8
		Civil	3	33.3
		Other	2	50
	Switch	Mechanical	0	0
		Civil	1	11.1
		Other	0	0
How do you feel about having the simulation try to correct your answers now that you know it may have compared your answers with an incorrect one?	Complacency	Mechanical	14	66.7
		Civil	6	66.7
		Other	3	75.0
	Vigilance	Mechanical	1	4.8
		Civil	3	33.3
		Other	0	0
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Vigilance	Mechanical	2	9.5
		Civil	0	0
		Other	0	0
	Switch	Mechanical	1	4.8
		Civil	0	0
		Other	0	0
How would you ensure that a guided simulation or other technology you are using isn't faulty?	Vigilance	Mechanical	12	57.1
		Civil	5	55.6
		Other	2	50.0

[a]Code was an emergent code added while analyzing the data

Appendix E

Tables relating to Automation Bias Qualitative Coding of Questionnaire and Interviews

This appendix contains the tables related to each of the Reliance on Technology questionnaire and the interviews that pertain to automation bias, separating out the grouping categories for individual reporting. Table 53 shows the frequency with which certain codes showed up in participant responses to the questions from the questionnaires and interviews divided into whether the participant had a correctly or an incorrectly working simulation. Table 54 shows the frequency with which certain codes showed up in participant responses to the questions from the questionnaires and interviews divided into which simulation group the participant belongs to. Table 55 shows the frequency with which certain codes showed up in participant responses to the questions from the questionnaires and interviews divided by participant confidence. Table 56 shows the frequency with which certain codes showed up in participant responses to the questions from the questionnaires and interviews divided by participant experience. Table 57 shows the frequency with which certain codes showed up in participant responses to the questions from the questionnaires and interviews divided by participant experience.

Qualitative Coding of Automation Bias based on Correctness of Simulation

Table 53

Qualitative Coding of Automation Bias, Simulation Correctness

Question	Coding	Group	Count	%
How do you feel about using guided simulations to help learn a topic?	Trust	Correct	12	70.6
		Incorrect	9	52.9
	Help	Correct	0	0
		Incorrect	2	11.8
Did you change your answer when given the opportunity by the simulation? If so, why?[b]	Trust	Correct	3	17.6
		Incorrect	2	11.8
How did you recheck your answer before resubmitting it?	Trust	Correct	16	94.1
		Incorrect	13	76.5
	Help[a]	Correct	1	5.9
		Incorrect	1	5.9
How often did you use the help provided by the simulation?	Help	Correct	12	70.6
		Incorrect	9	52.9
Did the help provided by the simulation help you solve the truss problem? If so, how?	Help	Correct	14	82.4
		Incorrect	5	29.4
	Trust	Correct	13	76.5
		Incorrect	4	23.5
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Awareness	Correct	4	23.5
		Incorrect	8	47.1
	Trust[a]	Correct	1	5.9
		Incorrect	2	11.8
What did you do if you disagreed with the simulation?	Awareness	Correct	0	0
		Incorrect	0	0
	Trust	Correct	2	11.8
		Incorrect	2	11.8
	Help[a]	Correct	1	5.9
		Incorrect	1	5.9
How do you feel about the fact that the purpose of the study was not shared from the beginning?	Trust	Correct	0	0
		Incorrect	1	5.9
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Awareness	Correct	0	0
		Incorrect	2	11.8
	Trust[a]	Correct	8	47.1
		Incorrect	3	17.6
How does knowing the information above about the intentional errors impact your thoughts on the simulation?	Trust	Correct	10	58.8
		Incorrect	6	35.3

Question	Coding	Group	Count	%
Has your perception of the help provided by the simulation changed after learning the true nature of the study and the simulation?	Help	Correct	4	23.5
		Incorrect	2	11.8
	Trust	Correct	4	23.5
		Incorrect	1	5.9
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Trust	Correct	8	47.1
		Incorrect	11	64.7
	Awareness	Correct	0	0
		Incorrect	1	5.9
How does this study impact your perceived usage of guided simulation moving forward?	Trust	Correct	5	29.4
		Incorrect	3	17.6
	Awareness	Correct	0	0
		Incorrect	0	0
	Help	Correct	1	5.9
		Incorrect	0	0
When is the ideal time to know that a guided simulation or other technology might be faulty?	Awareness	Correct	0	0
		Incorrect	0	0
How would you ensure that a guided simulation or other technology you are using isn't faulty?[b]	Help[a]	Correct	2	11.8
		Incorrect	3	17.6
Tell me about your experience using the guided simulation.	Trust	Correct	0	0
		Incorrect	2	66.7
	Awareness	Correct	0	0
		Incorrect	1	33.3
What was your strategy in solving the truss problem?	Trust	Correct	1	100
		Incorrect	0	0
	Help[a]	Correct	1	100
		Incorrect	0	0
How often did the simulation disagree with your answer?[b]	Awareness	Correct	0	0
		Incorrect	2	66.7
When the program disagreed with your answer, what did you do?	Awareness	Correct	0	0
		Incorrect	1	33.3
	Trust[a]	Correct	0	0
		Incorrect	2	66.7
What was your experience using the help provided by the simulation?	Help	Correct	1	100
		Incorrect	1	33.3

Question	Coding	Group	Count	%
How did you feel when you found out the guided simulation might intentionally contain errors?	Awareness	Correct	1	100
		Incorrect	3	100
	Trust	Correct	1	100
		Incorrect	1	33.3

[a]Code was an emergent code added while analyzing the data. [b]Question received emergent code from an a priori automation bias code despite being an automation complacency question.

Qualitative Coding of Automation Bias based on Simulation Group

Table 54

Qualitative Coding of Automation Bias, Simulation Group

Question	Coding	Group	Count	%
How do you feel about using guided simulations to help learn a topic?	Trust	Control	8	88.9
		Deceived	5	55.6
		Informed	4	50.0
		Misled	4	50.0
	Help	Control	0	0
		Deceived	0	0
		Informed	0	0
		Misled	2	25.0
Did you change your answer when given the opportunity by the simulation? If so, why?[b]	Trust	Control	2	22.2
		Deceived	1	11.1
		Informed	1	12.5
		Misled	1	12.5
How did you recheck your answer before resubmitting it?	Trust	Control	9	100
		Deceived	8	88.9
		Informed	7	87.5
		Misled	5	62.5
	Help[a]	Control	1	11.1
		Deceived	1	11.1
		Informed	0	0
		Misled	0	0
How often did you use the help provided by the simulation?	Help	Control	5	55.6
		Deceived	6	66.7
		Informed	7	87.5
		Misled	3	37.5
Did the help provided by the simulation help you solve the truss problem? If so, how?	Help	Control	6	66.7
		Deceived	4	44.4
		Informed	8	100
		Misled	1	12.5
	Trust	Control	5	55.6
		Deceived	3	33.3
		Informed	8	100
		Misled	1	12.5

Question	Coding	Group	Count	%
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Awareness	Control	2	22.2
		Deceived	4	44.4
		Informed	2	25.0
		Misled	4	50.0
	Trust[a]	Control	0	0
		Deceived	2	22.2
		Informed	1	12.5
		Misled	0	0
What did you do if you disagreed with the simulation?	Awareness	Control	0	0
		Deceived	0	0
		Informed	0	0
		Misled	0	0
	Trust	Control	1	11.1
		Deceived	1	11.1
		Informed	1	12.5
		Misled	1	12.5
	Help[a]	Control	1	11.1
		Deceived	1	11.1
		Informed	0	0
		Misled	0	0
How do you feel about the fact that the purpose of the study was not shared from the beginning?	Trust	Control	0	0
		Deceived	1	11.1
		Informed	0	0
		Misled	0	0
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Awareness	Control	0	0
		Deceived	1	11.1
		Informed	0	0
		Misled	1	11.1
	Trust[a]	Control	5	55.6
		Deceived	2	22.2
		Informed	3	37.5
		Misled	1	12.5
How does knowing the information above about the intentional errors impact your thoughts on the simulation?	Trust	Control	6	66.7
		Deceived	3	33.3
		Informed	4	50.0
		Misled	3	37.5
Has your perception of the help provided by the simulation changed after learning the true nature of the study and the simulation?	Help	Control	2	22.2
		Deceived	1	11.1
		Informed	2	25.0
		Misled	1	12.5
	Trust	Control	2	22.2
		Deceived	0	0
		Informed	2	25.0
		Misled	1	12.5

Question	Coding	Group	Count	%
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Trust	Control	3	33.3
		Deceived	6	66.7
		Informed	5	62.5
		Misled	5	62.5
	Awareness	Control	0	0
		Deceived	0	0
		Informed	0	0
		Misled	1	12.5
How does this study impact your perceived usage of guided simulation moving forward?	Trust	Control	2	22.2
		Deceived	2	22.2
		Informed	3	37.5
		Misled	1	12.5
	Awareness	Control	0	0
		Deceived	0	0
		Informed	0	0
		Misled	0	0
	Help	Control	0	0
		Deceived	0	0
		Informed	1	12.5
		Misled	0	0
When is the ideal time to know that a guided simulation or other technology might be faulty?	Awareness	Control	0	0
		Deceived	0	0
		Informed	0	0
		Misled	0	0
How would you ensure that a guided simulation or other technology you are using isn't faulty?[b]	Help[a]	Control	0	0
		Deceived	2	22.2
		Informed	2	25.0
		Misled	1	12.5
Tell me about your experience using the guided simulation.	Trust	Control	0	0
		Deceived	0	0
		Informed	0	0
		Misled	2	100
	Awareness	Control	0	0
		Deceived	0	0
		Informed	0	0
		Misled	1	50.0
What was your strategy in solving the truss problem?	Trust	Control	1	100
		Deceived	0	0
		Informed	0	0
		Misled	0	0
	Help[a]	Control	1	100
		Deceived	0	0
		Informed	0	0
		Misled	0	0
	Awareness	Control	0	0%

Question	Coding	Group	Count	%
How often did the simulation disagree with your answer?[b]		Deceived	0	0
		Informed	0	0
		Misled	2	100
When the program disagreed with your answer, what did you do?	Awareness	Control	0	0
		Deceived	0	0
		Informed	0	0
		Misled	1	50.0
	Trust[a]	Control	0	0
		Deceived	1	100
		Informed	0	0
		Misled	1	50.0
What was your experience using the help provided by the simulation?	Help	Control	1	100
		Deceived	1	100
		Informed	0	0
		Misled	0	0
How did you feel when you found out the guided simulation might intentionally contain errors?	Awareness	Control	1	100
		Deceived	1	100
		Informed	0	0
		Misled	2	100
	Trust	Control	1	100
		Deceived	1	100
		Informed	0	0
		Misled	0	0

[a]Code was an emergent code added while analyzing the data. [b]Question received emergent code from an a priori automation bias code despite being an automation complacency question.

Qualitative Coding of Automation Bias based on Confidence

Table 55

Qualitative Coding of Automation Bias, Confidence

Question	Coding	Confidence	Count	Percentage
How do you feel about using guided simulations to help learn a topic?	Trust	High	19	61.3%
		Low	2	66.7%
	Help	High	2	6.5%
		Low	0	0%
Did you change your answer when given the opportunity by the simulation? If so, why?[b]	Trust[a]	High	5	16.1%
		Low	0	0%
How did you recheck your answer before resubmitting it?	Trust	High	27	87.1%
		Low	2	66.7%
	Help[a]	High	2	6.5%
		Low	0	0%
How often did you use the help provided by the simulation?	Help	High	20	64.5%
		Low	1	33.3%
Did the help provided by the simulation help you solve the truss problem? If so, how?	Help	High	18	58.1%
		Low	1	33.3%
	Trust	High	16	51.6%
		Low	1	33.3%
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Awareness	High	12	38.7%
		Low	0	0%
	Trust[a]	High	3	9.7%
		Low	0	0%
What did you do if you disagreed with the simulation?	Awareness	High	0	0%
		Low	0	0%
	Trust	High	4	12.9%
		Low	0	0%
	Help[a]	High	2	6.5%
		Low	0	0%
How do you feel about the fact that the purpose of the study was not shared from the beginning?	Trust	High	1	3.2%
		Low	0	0%
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Awareness	High	2	6.5%
		Low	0	0%
	Trust[a]	High	10	32.3%
		Low	1	33.3%
How does knowing the information above about the intentional errors impact your thoughts on the simulation?	Trust	High	15	48.4%
		Low	1	33.3%

Question	Coding	Confidence	Count	Percentage
Has your perception of the help provided by the simulation changed after learning the true nature of the study and the simulation?	Help	High	5	16.1%
		Low	1	33.3%
	Trust	High	4	12.9%
		Low	1	33.3%
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Trust	High	17	54.8%
		Low	2	66.7%
	Awareness	High	1	3.2%
		Low	0	0%
How does this study impact your perceived usage of guided simulation moving forward?	Trust	High	8	25.8%
		Low	0	0%
	Awareness	High	0	0%
		Low	0	0%
	Help	High	1	3.2%
		Low	0	0%
When is the ideal time to know that a guided simulation or other technology might be faulty?	Awareness	High	0	0%
		Low	0	0%
How would you ensure that a guided simulation or other technology you are using isn't faulty?[b]	Help[a]	High	5	16.1%
		Low	0	0%
Tell me about your experience using the guided simulation.	Trust	High	2	50%
		Low	0	0%
	Awareness	High	1	25%
		Low	0	0%
What was your strategy in solving the truss problem?	Trust	High	1	25%
		Low	0	0%
	Help[a]	High	1	25%
		Low	0	0%
How often did the simulation disagree with your answer?[b]	Awareness	High	2	50%
		Low	0	0%
When the program disagreed with your answer, what did you do?	Awareness	High	1	25%
		Low	0	0%
	Trust[a]	High	2	50%
		Low	0	0%
What was your experience using the help provided by the simulation?	Help	High	2	50%
		Low	0	0%

Question	Coding	Confidence	Count	Percentage
How did you feel when you found out the guided simulation might intentionally contain errors?	Awareness	High	4	100%
		Low	0	0%
	Trust	High	2	50%
		Low	0	0%

[a]Code was an emergent code added while analyzing the data. [b]Question received emergent code from an a priori automation bias code despite being an automation complacency question.

Qualitative Coding of Automation Bias based on Trust

Table 56

Qualitative Coding of Automation Bias, Trust

Question	Coding	Trust	Count	Percentage
How do you feel about using guided simulations to help learn a topic?	Trust	High	13	65%
		Neutral	3	100%
		Low	5	45.5%
	Help	High	1	5%
		Neutral	0	0%
		Low	1	9.1%
Did you change your answer when given the opportunity by the simulation? If so, why?[b]	Trust	High	4	20%
		Neutral	0	0%
		Low	1	9.1%
How did you recheck your answer before resubmitting it?	Trust	High	19	95%
		Neutral	3	100%
		Low	7	63.6%
	Help[a]	High	1	5%
		Neutral	0	0%
		Low	1	9.1%
How often did you use the help provided by the simulation?	Help	High	14	70%
		Neutral	3	100%
		Low	4	36.4%
Did the help provided by the simulation help you solve the truss problem? If so, how?	Help	High	15	75%
		Neutral	3	100%
		Low	1	9.1%
	Trust	High	14	70%
		Neutral	2	66.7%
		Low	1	9.1%
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Awareness	High	4	20%
		Neutral	2	66.7%
		Low	6	54.5%
	Trust[a]	High	2	10%
		Neutral	1	33.3%
		Low	0	0%
What did you do if you disagreed with the simulation?	Awareness	High	0	0%
		Neutral	0	0%
		Low	0	0%
	Trust	High	4	20%
		Neutral	0	0%
		Low	0	0%
	Help[a]	High	1	5%
		Neutral	1	33.3%
		Low	0	0%

Question	Coding	Trust	Count	Percentage
How do you feel about the fact that the purpose of the study was not shared from the beginning?	Trust	High	1	5%
		Neutral	0	0%
		Low	0	0%
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Awareness	High	0	0%
		Neutral	1	33.3%
		Low	1	9.1%
	Trust[a]	High	10	50%
		Neutral	0	0%
		Low	1	9.1%
How does knowing the information above about the intentional errors impact your thoughts on the simulation?	Trust	High	12	60%
		Neutral	2	66.7%
		Low	2	18.2%
Has your perception of the help provided by the simulation changed after learning the true nature of the study and the simulation?	Help	High	4	20%
		Neutral	0	0%
		Low	2	18.2%
	Trust	High	4	20%
		Neutral	0	0%
		Low	1	9.1%
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Trust	High	9	45%
		Neutral	2	66.7%
		Low	8	72.7%
	Awareness	High	1	5%
		Neutral	0	0%
		Low	0	0%
How does this study impact your perceived usage of guided simulation moving forward?	Trust	High	5	25%
		Neutral	2	66.7%
		Low	1	9.1%
	Awareness	High	0	0%
		Neutral	0	0%
		Low	0	0%
	Help	High	1	5%
		Neutral	0	0%
		Low	0	0%
When is the ideal time to know that a guided simulation or other technology might be faulty?	Awareness	High	0	0%
		Neutral	0	0%
		Low	0	0%
How would you ensure that a guided simulation or other technology you are using isn't faulty?[b]	Help[a]	High	4	20%
		Neutral	1	33.3%
		Low	0	0%

Question	Coding	Trust	Count	Percentage
Tell me about your experience using the guided simulation.	Trust	High	0	0%
		Neutral	0	0%
		Low	2	100%
	Awareness	High	0	0%
		Neutral	0	0%
		Low	1	50%
What was your strategy in solving the truss problem?	Trust	High	1	100%
		Neutral	0	0%
		Low	0	0%
	Help[a]	High	1	100%
		Neutral	0	0%
		Low	0	0%
How often did the simulation disagree with your answer?[b]	Awareness	High	0	0%
		Neutral	0	0%
		Low	2	100%
When the program disagreed with your answer, what did you do?	Awareness	High	0	0%
		Neutral	0	0%
		Low	1	50%
	Trust[a]	High	0	0%
		Neutral	1	100%
		Low	1	50%
What was your experience using the help provided by the simulation?	Help	High	1	100%
		Neutral	1	100%
		Low	0	0%
How did you feel when you found out the guided simulation might intentionally contain errors?	Awareness	High	1	100%
		Neutral	1	100%
		Low	2	100%
	Trust	High	1	100%
		Neutral	1	100%
		Low	0	0%

[a]Code was an emergent code added while analyzing the data. [b]Question received emergent code from an a priori automation bias code despite being an automation complacency question.

Qualitative Coding of Automation Bias based on Experience

Table 57

Qualitative Coding of Automation Bias, Experience

Question	Coding	Experience	Count	Percentage
How do you feel about using guided simulations to help learn a topic?	Trust	Very High	2	50%
		High	1	16.7%
		Low	18	75%
	Help	Very High	0	0%
		High	0	0%
		Low	2	8.3%
Did you change your answer when given the opportunity by the simulation? If so, why?[b]	Trust	Very High	0	0%
		High	2	33.3%
		Low	3	12.5%
How did you recheck your answer before resubmitting it?	Trust	Very High	3	75%
		High	4	66.7%
		Low	22	91.7%
	Help[a]	Very High	0	0%
		High	0	0%
		Low	2	8.3%
How often did you use the help provided by the simulation?	Help	Very High	2	50%
		High	3	50%
		Low	16	66.7%
Did the help provided by the simulation help you solve the truss problem? If so, how?	Help	Very High	2	50%
		High	4	66.7%
		Low	13	54.2%
	Trust	Very High	2	50%
		High	4	66.7%
		Low	11	45.8%
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Awareness	Very High	0	0%
		High	1	16.7%
		Low	11	45.8%
	Trust[a]	Very High	1	25%
		High	0	0%
		Low	2	8.3%
What did you do if you disagreed with the simulation?	Awareness	Very High	0	0%
		High	0	0%
		Low	0	0%
	Trust	Very High	1	25%
		High	0	0%
		Low	3	12.5%
	Help[a]	Very High	0	0%
		High	1	16.7%
		Low	1	4.2%

Question	Coding	Experience	Count	Percentage
How do you feel about the fact that the purpose of the study was not shared from the beginning?	Trust	Very High	0	0%
		High	0	0%
		Low	1	4.2%
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Awareness	Very High	0	0%
		High	0	0%
		Low	2	12.5%
	Trust[a]	Very High	1	25%
		High	2	33.3%
		Low	8	33.3%
How does knowing the information above about the intentional errors impact your thoughts on the simulation?	Trust	Very High	3	75%
		High	3	50%
		Low	10	41.7%
Has your perception of the help provided by the simulation changed after learning the true nature of the study and the simulation?	Help	Very High	0	0%
		High	2	33.3%
		Low	4	16.7%
	Trust	Very High	0	0%
		High	2	33.3%
		Low	3	12.5%
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Trust	Very High	2	50%
		High	2	33.3%
		Low	15	62.5%
	Awareness	Very High	0	0%
		High	0	0%
		Low	1	4.2%
How does this study impact your perceived usage of guided simulation moving forward?	Trust	Very High	0	0%
		High	3	50%
		Low	5	20.8%
	Awareness	Very High	0	0%
		High	0	0%
		Low	0	0%
	Help	Very High	1	25%
		High	0	0%
		Low	0	0%
When is the ideal time to know that a guided simulation or other technology might be faulty?	Awareness	Very High	0	0%
		High	0	0%
		Low	0	0%

Question	Coding	Experience	Count	Percentage
How would you ensure that a guided simulation or other technology you are using isn't faulty?[b]	Help[a]	Very High	1	25%
		High	0	0%
		Low	4	16.7%

[a]Code was an emergent code added while analyzing the data. [b]Question received emergent code from an a priori automation bias code despite being an automation complacency question.

Qualitative Coding of Automation Bias based on Gender

Table 58

Qualitative Coding of Automation Bias, Gender

Question	Coding	Gender	Count	%
How do you feel about using guided simulations to help learn a topic?	Trust	Male	14	56.0
		Female	7	77.8
	Help	Male	1	4.0
		Female	1	11.1
Did you change your answer when given the opportunity by the simulation? If so, why?[b]	Trust	Male	3	12.0
		Female	2	22.2
How did you recheck your answer before resubmitting it?	Trust	Male	21	84.0
		Female	8	88.9
	Help[a]	Male	0	0
		Female	2	22.2
How often did you use the help provided by the simulation?	Help	Male	14	56.0
		Female	7	77.8
Did the help provided by the simulation help you solve the truss problem? If so, how?	Help	Male	13	52.0
		Female	6	66.7
	Trust	Male	12	48.0
		Female	5	55.6
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Awareness	Male	8	72.0
		Female	4	44.4
	Trust[a]	Male	3	12.0
		Female	0	0
What did you do if you disagreed with the simulation?	Awareness	Male	0	0
		Female	0	0
	Trust	Male	4	16.0
		Female	0	0
	Help[a]	Male	2	8.0
		Female	0	0
How do you feel about the fact that the purpose of the study was not shared from the beginning?	Trust	Male	1	4.0
		Female	0	0
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Awareness	Male	2	8.0
		Female	0	0
	Trust[a]	Male	8	32.0
		Female	3	33.3
How does knowing the information above about the intentional errors impact your thoughts on the simulation?	Trust	Male	15	60.0
		Female	1	11.1

Question	Coding	Gender	Count	%
Has your perception of the help provided by the simulation changed after learning the true nature of the study and the simulation?	Help	Male	5	20.0
		Female	1	11.1
	Trust	Male	4	16.0
		Female	1	11.1
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Trust	Male	13	52.0
		Female	6	66.7
	Awareness	Male	0	0
		Female	1	11.1
How does this study impact your perceived usage of guided simulation moving forward?	Trust	Male	6	24.0
		Female	2	22.2
	Awareness	Male	0	0
		Female	0	0
	Help	Male	0	0
		Female	1	11.1
When is the ideal time to know that a guided simulation or other technology might be faulty?	Awareness	Male	0	0
		Female	0	0
How would you ensure that a guided simulation or other technology you are using isn't faulty?[b]	Help[a]	Male	3	12.0
		Female	2	22.2

[a]Code was an emergent code added while analyzing the data. [b]Question received emergent code from an a priori automation bias code despite being an automation complacency question.

Qualitative Coding of Automation Bias based on Age

Table 59

Qualitative Coding of Automation Bias, Age

Question	Coding	Age	Count	%
How do you feel about using guided simulations to help learn a topic?	Trust	18-21	10	55.6
		22-24	8	61.5
		25+	3	100
	Help	18-21	1	5.6
		22-24	1	7.7
		25+	0	0
Did you change your answer when given the opportunity by the simulation? If so, why?[b]	Trust	18-21	3	16.7
		22-24	2	15.4
		25+	0	0
How did you recheck your answer before resubmitting it?	Trust	18-21	14	77.8
		22-24	12	92.3
		25+	3	100
	Help[a]	18-21	2	11.1
		22-24	0	0
		25+	0	0
How often did you use the help provided by the simulation?	Help	18-21	11	61.1
		22-24	9	69.2
		25+	1	33.3
Did the help provided by the simulation help you solve the truss problem? If so, how?	Help	18-21	10	55.6
		22-24	8	61.5
		25+	1	33.3
	Trust	18-21	9	50.0
		22-24	7	53.8
		25+	1	33.3
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Awareness	18-21	6	33.3
		22-24	5	38.5
		25+	1	33.3
	Trust[a]	18-21	3	16.7
		22-24	0	0
		25+	0	0
What did you do if you disagreed with the simulation?	Awareness	18-21	0	0
		22-24	0	0
		25+	0	0
	Trust	18-21	2	11.1
		22-24	1	7.7
		25+	1	33.3
	Help[a]	18-21	1	5.6
		22-24	1	7.7
		25+	0	0

Question	Coding	Age	Count	%
How do you feel about the fact that the purpose of the study was not shared from the beginning?	Trust	18-21	1	5.6
		22-24	0	0
		25+	0	0
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Awareness	18-21	1	5.6
		22-24	1	7.7
		25+	0	0
	Trust[a]	18-21	4	22.2
		22-24	5	38.5
		25+	2	66.7
How does knowing the information above about the intentional errors impact your thoughts on the simulation?	Trust	18-21	6	33.3
		22-24	8	61.5
		25+	2	66.7
Has your perception of the help provided by the simulation changed after learning the true nature of the study and the simulation?	Help	18-21	3	16.7
		22-24	3	23.1
		25+	0	0
	Trust	18-21	2	11.1
		22-24	3	23.1
		25+	0	0
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Trust	18-21	11	61.1
		22-24	7	5.8
		25+	1	33.3
	Awareness	18-21	1	5.6
		22-24	0	0
		25+	0	0
How does this study impact your perceived usage of guided simulation moving forward?	Trust	18-21	4	22.2
		22-24	4	30.8
		25+	0	0
	Awareness	18-21	0	0
		22-24	0	0
		25+	0	0
	Help	18-21	1	5.6
		22-24	0	0
		25+	0	0
When is the ideal time to know that a guided simulation or other technology might be faulty?	Awareness	18-21	0	0
		22-24	0	0
		25+	0	0

Question	Coding	Age	Count	%
How would you ensure that a guided simulation or other technology you are using isn't faulty?[b]	Help[a]	18-21	4	22.2
		22-24	0	0
		25+	1	33.3

[a]Code was an emergent code added while analyzing the data. [b]Question received emergent code from an a priori automation bias code despite being an automation complacency question.

Qualitative Coding of Automation Bias based on Major

Table 60

Qualitative Coding of Automation Bias, Major

Question	Coding	Major	Count	Percentage
How do you feel about using guided simulations to help learn a topic?	Trust	Mechanical	11	52.4
		Civil	7	77.8
		Other	3	75.0
	Help	Mechanical	1	4.8
		Civil	1	11.1
		Other	0	0
Did you change your answer when given the opportunity by the simulation? If so, why?[b]	Trust	Mechanical	3	14.3
		Civil	0	0
		Other	2	50.0
How did you recheck your answer before resubmitting it?	Trust	Mechanical	17	81.0
		Civil	8	88.9
		Other	4	100
	Help[a]	Mechanical	2	9.5
		Civil	0	0
		Other	0	0
How often did you use the help provided by the simulation?	Help	Mechanical	14	66.7
		Civil	5	55.6
		Other	2	50.0
Did the help provided by the simulation help you solve the truss problem? If so, how?	Help	Mechanical	12	57.1
		Civil	5	55.6
		Other	2	50.0
	Trust	Mechanical	11	52.4
		Civil	4	44.4
		Other	2	50.0
Did the simulation tell you your answer was wrong at any point? Did you agree with the simulation? Why or why not?	Awareness	Mechanical	8	38.1
		Civil	2	22.2
		Other	2	50.0
	Trust[a]	Mechanical	3	14.3
		Civil	0	0
		Other	0	0
What did you do if you disagreed with the simulation?	Awareness	Mechanical	0	0
		Civil	0	0
		Other	0	0
	Trust	Mechanical	3	14.3
		Civil	0	0
		Other	1	11.1
	Help[a]	Mechanical	0	0
		Civil	1	11.1
		Other	1	25.0

Question	Coding	Major	Count	Percentage
How do you feel about the fact that the purpose of the study was not shared from the beginning?	Trust	Mechanical	1	4.8
		Civil	0	0
		Other	0	0
Did you suspect there was anything wrong with the guided simulation? If so, when and what did you suspect?	Awareness	Mechanical	2	9.5
		Civil	0	0
		Other	0	0
	Trust[a]	Mechanical	6	28.6
		Civil	3	33.3
		Other	2	50.0
How does knowing the information above about the intentional errors impact your thoughts on the simulation?	Trust	Mechanical	9	42.9
		Civil	5	55.6
		Other	2	50.0
Has your perception of the help provided by the simulation changed after learning the true nature of the study and the simulation?	Help	Mechanical	3	14.3
		Civil	2	22.2
		Other	1	25.0
	Trust	Mechanical	2	9.5
		Civil	2	22.2
		Other	1	25.0
What are your thoughts about the process now that you know your simulation {did/not} contain intentional errors?	Trust	Mechanical	12	57.1
		Civil	5	55.6
		Other	2	50.0
	Awareness	Mechanical	1	4.8
		Civil	0	0
		Other	0	0
How does this study impact your perceived usage of guided simulation moving forward?	Trust	Mechanical	5	23.8
		Civil	0	0
		Other	3	75.0
	Awareness	Mechanical	0	0
		Civil	0	0
		Other	0	0
	Help	Mechanical	1	4.8
		Civil	0	0
		Other	0	0
When is the ideal time to know that a guided simulation or other technology might be faulty?	Awareness	Mechanical	0	0
		Civil	0	0
		Other	0	0

Question	Coding	Major	Count	Percentage
How would you ensure that a guided simulation or other technology you are using isn't faulty?[b]	Help[a]	Mechanical	5	23.8
		Civil	1	11.1
		Other	0	0

[a]Code was an emergent code added while analyzing the data. [b]Question received emergent code from an a priori automation bias code despite being an automation complacency question.

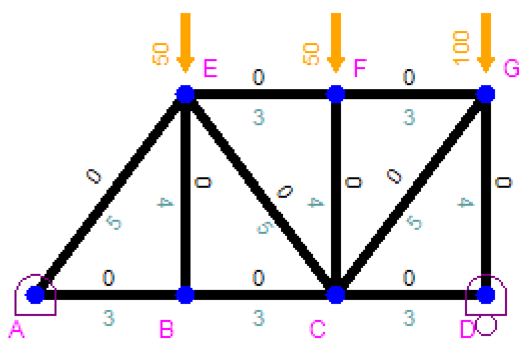
Appendix F

Method of Joints

Extra Credit Assignment

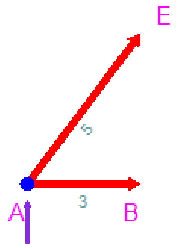
For this assignment, you will work through a truss system and solve it using the Method of Joints.

Here is the Truss System. The external forces are Orange with their forces listed. The green numbers are the lengths of the members, while the purple letters are the truss letters. The forces are currently listed as 0, as they hare unfound. As you work through the truss, the colors will change to show whether the member is in tension (red) or compression (green). You can use this to check your work as you go.

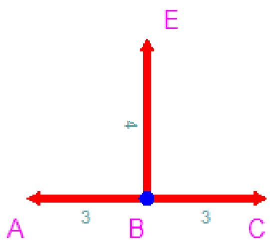


1) Solve the reaction forces generated by the support.

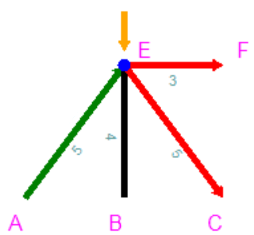
2) Solve the members attached to Joint A



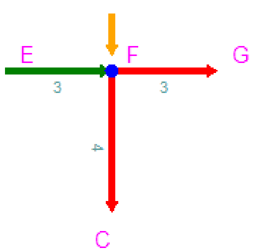
3) Solve the members attached to Joint B



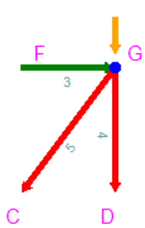
4) Solve the members attached to joint E



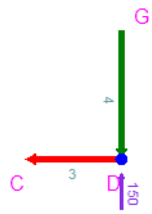
5) Solve the members attached to joint F



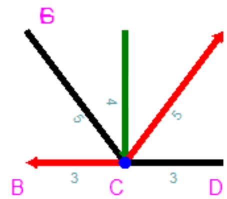
6) Solve the members attached to joint G



7) Solve the members attached to joint D



8) Confirm that the truss solution works



Appendix G

Informed Consent

Impact of Guided Simulation on Learning

Introduction

You are invited to participate in a research study conducted by Jon Anderson, a graduate student in the Engineering Education Department at Utah State University. This study is being supervised by the Principal Investigator, Wade Goodridge, a faculty member in the Engineering Education Department at Utah State University. The purpose of this research is to provide insight into how students interact with a computer-based simulation while attempting to learn a method of truss analysis in a Statics class. Your participation is entirely voluntary.

Procedures

Your participation will involve completing a demographic/previous experience survey, access to a software program designed to help you work through a truss analysis problem using the method of joints, a post survey, and potentially an interview about your experience using the software. The surveys are intended to take between 5 to 30 minutes each, while working in the program is intended to take roughly an hour of time to work through the truss problem with the computer aid. Eight participants will be invited, based on their work in the simulation and the surveys, to participate in an interview, and the interview should take between 30 minutes to 1 hour. Your total participation in this project is expected to be between one and a half and two and a half hours. If you agree to participate, all information related to the study will be collected through the surveys, including submitting the final results from the computer simulation to the final survey. We anticipate that 80 to 120 people will participate in this research study.

Alternative Procedures

Rather than participate in this research, you might prefer alternatives such as access to a guided notes assignment in place of the software package.

Risks

This is a minimal risk research study. That means that the risks of participating are no more likely or serious than those you encounter in everyday activities. The foreseeable risks or discomforts include a potential loss of confidentiality if records are accidentally made available. This risk is highest for those participants who are interviewed. In order to minimize those risks and discomforts, the researchers will anonymize the data and strictly maintain a separation of the immediate results of individuals from the identifying information. In addition, students interviewed will be given pseudonyms and any identifying remarks made in the interview will be redacted or adjusted to remove the identifiable portions of the conversation. If you have a bad research-related experience, please contact Jon Anderson.

Benefits

Participation in this study may directly benefit you by exposing you to an intervention that can help you better understand the topic of the method of joints. We cannot guarantee that you will directly benefit from this study, but it has been designed to learn more about the way in which students interact with software study aids. This should greatly increase our understanding and allow for better computer simulations to be made in the future.

Confidentiality

The researchers will make every effort to ensure that the information you provide as part of this study remains confidential. Your identity will not be revealed in any publications, presentations, or reports resulting from this research study.

We will collect your information through Qualtrics surveys, your computer guided simulation file upload, and video recordings of those selected for interviews. Online activities always carry a risk of a data breach, but we will use systems and processes that minimize breach opportunities. This information will be securely stored in a restricted-access folder on Box.com, an encrypted, cloud-based storage system. Links between the study generated identifier and the actual identification will be kept in separate folders on Box.com to ensure that they are not linked easily. The actual identification will be destroyed three months after the paper is published to ensure that your identity is protected. Video recordings will be destroyed at the same time, although the transcription of said recordings will be retained. This form/survey, identifying your acknowledgement of the informed consent, will be kept for three years after the study is complete, and then it will be destroyed.

It is unlikely, but possible, that others (Utah State University or state or federal officials) may require us to share the information you give us from the study to ensure that the research was conducted safely and appropriately. We will only share your information if law or policy requires us to do so.

Voluntary Participation & Withdrawal

Your participation in this research is completely voluntary. If you agree to participate now and change your mind later, you may withdraw at any time by contacting Jon Anderson at A02231235@usu.edu. If you choose to withdraw after we have already collected information about you, we will remove your information from all of our data stores, including your results from our Qualtrics surveys, the results of your computer simulation, and your interviews, as applicable.

The researchers may choose to terminate your participation in this research study if they find that you are under 18 years of age, or if you fail to complete all aspects of the study, including either survey or the file provided at the end of the computer simulation. In addition, simply selecting the same response for all answers to the survey questions will have your results removed from the study. Finally, this study is intended to determine how well each individual does with the simulation; if we discover that you have collaborated with others while working on this study, both parties will be removed from the study.

Compensation

For your participation in this research study, you will receive 10 points of extra credit. You will receive this compensation upon completing the final survey, following your use of the computer simulation. As an alternative, you may choose to complete a worksheet covering the same type of problem to earn this extra credit, but your results will not be part of the study. If you are removed from the study for any reason listed above, you may still complete this alternative worksheet to receive the credit.

You should be aware while considering participation that your instructor will not be made aware of your performance on the simulation or the alternative assignment. They will also not be made aware of whether you participated in the survey or if you did the alternative assignment. All they will be informed of is that you participated by virtue of receiving the extra credit.

Findings

If the researchers learn anything new during the course of this research study that might affect your willingness to continue participation, you will be contacted about those findings. This might include changes in procedures, changes in the risks or benefits of participation, or any new alternatives to participation that the researchers learn about.

IRB Review

The Institutional Review Board (IRB) for the protection of human research participants at Utah State University has reviewed and approved this study. If you have questions about the research study itself, please contact the Principal

Investigator at (435) 797-9051 or wade.goodridge@usu.edu. If you have questions about your rights or would simply like to speak with someone *other* than the research team about questions or concerns, please contact the IRB Director at (435) 797-0567 or irb@usu.edu.

Wade Goodridge	*Jon Anderson*
Wade Goodridge Principal Investigator (435) 797-9051; wade.goodridge@usu.edu	Jon Anderson Student Investigator (801) 863-6575; A02231235@usu.edu

Informed Consent

By signing below, you agree to participate in this study. You indicate that you understand the risks and benefits of participation, and that you know what you will be asked to do. You also agree that you have asked any questions you might have, and are clear on how to stop your participation in the study if you choose to do so. Please be sure to retain a copy of this form for your records.

[Due to being a Qualtrics survey, the following radio button options will be given at the start of the survey:]

- I agree to participate in this study. I will complete the required information on the next question, including my name, date, and signature, before continuing on to the rest of the demographic data.
- I do not agree to participate in this study. I recognize that this means that this survey will now end, and no additional information will be provided.

Participant's Signature	Participant's Name, Printed	Date

Appendix H

Qualitative Codebook

Table 61.

Qualitative Codebook used to code Free Response Questions in the Reliance on Technology Questionnaire and the Interviews

Code	Definition	Purpose/meaning of code	In vivo description
Trust	The participant trusted the simulation.	Measures whether the participant is using following the advice of the simulation to solve the provided joints. *Did the participant comment on following the advice of the simulation?*	"In general, I think that guided simulations can be very beneficial in helping students visualize what they are working on. I believe that this simulation did a very good job of showing the user what was going on in the truss, and helped me to solve it faster."
Awareness	The participant is aware of the state of the system; i.e., if it has errors or not.	Measures whether the participant is aware of which simulation they have; is the simulation correct or faulty? *Did the participant indicate that they knew something was wrong with the simulation?*	"The simulation would tell me my answer was wrong and I did not agree with it. With what the simulation had explained about tension and compression it was confusing because they would sometimes flip a sign that I thought was right."
Help	The participant sought help during the procedure.	Measures whether the participant clicked on the help button provided by the simulation. *Did the participant ever press the help button?*	"I would use the help provided when I got the answer wrong the first time."

Code	Definition	Purpose/meaning of code	In vivo description
Complacency	The participant checked their answers when told by the simulation.	Measures whether the participants are following the directions of the simulation, particularly when they get an answer wrong. *Did the participant check their answer and resubmit a new value?*	"Quite a few times, mostly because I need to refresh on tension and compression."
Vigilance	The participant is actively checking to ensure the simulation is correct	Measures whether the participant indicates that they are checking the simulation itself. *Did the participant indicate that they were aware of the correct state of the simulation?*	"No, I didn't. The dimensions were different in the FBD than in the given truss, so I gave up and figured something was wrong with the values that I was using or the values the program was using."
Switch	The participant changed their answer	Measures whether the participant recognizes that they changed their answer because the simulation told them to. *Did the participant acknowledge that they changed any of their answers?*	"Yes, being prompted that I was wrong made me rethink the process."

CURRICULUM VITA

Jon Anderson, currently pursuing Ph.D.
1076 N 1700 W, Pleasant Grove, UT 84604
801-592-4935
jonathana@uvu.edu

EDUCATION

Utah State University, Logan, UT — **Currently Pursuing, Expected 2023**
Ph.D. in Engineering Education
ABD, Dissertation: "Student Reliance on Simulations: The Extent that Engineering Students Rely on the Outcomes of Their Simulations"

Brigham Young University, Provo, UT — **2007**
M.S. in Electrical and Computer Engineering
Thesis: "Semi Autonomous Vehicle Intelligence: Real Time Target Tracking for Vision Guided Autonomous Vehicles"

Brigham Young University, Provo, UT — **2004**
B.S. in Electrical and Computer Engineering

Utah Valley State College, Orem, UT — **2001**
A.S. in Science, with Honors

AWARDS AND HONORS

Civic Thought & Leadership Grant	**2021**
Alan E. Hall Innovation Award	**2015**
Dean's Faculty Scholarship Award	**2013**
Faculty Appreciation Award, Utah Valley University Math Lab	**2013**
Outstanding Teaching Award – Live Interactive	**2013**
SPIE Intelligent Robot and Computer Vision Conference Best Student Paper	**2005**

TEACHING EXPERIENCE

Utah Valley University — **2017-Present**
Associate Professor
Developed syllabus and overall course structure, wrote and graded exams, developed online course material for students, prepared additional lecture notes to aid students in difficult concepts, and administered grades.

Utah Valley University — **2013-2017**
Assistant Professor
Developed syllabus and overall course structure, wrote and graded exams, developed online course material for students, prepared additional lecture notes to aid students in difficult concepts, and administered grades.

Utah Valley University — **2011-2013**
Instructor
Developed syllabus and overall course structure, wrote and graded exams, developed online course material for students, prepared additional lecture notes to aid students in difficult concepts, and administered grades.

Utah Valley University — **2010-2011**
Lecturer
Developed syllabus and overall course structure, wrote and graded exams, developed online course material for students, prepared additional lecture notes to aid students in difficult concepts, and administered grades.

Utah Valley University — **2008-2010**
Adjunct Professor
Used departmental syllabus and overall course structure, wrote and graded exams, developed online course material for students, prepared additional lecture notes to aid students in difficult concepts, and administered grades.

Brigham Young University, Provo, UT **2002-2007**
Teaching Assistant – in the Electrical and Computer Engineering department. Collaborated on curriculum and exam development, met with students upon request, and graded all written work, including final exam papers.
Brigham Young University, Provo, UT **2005-2007**
Research Assistant – in the Electrical and Computer Engineering department. Worked with cameras, integrated circuits, and FPGAs to develop computer-aided vision guidance systems for remote control cars.

COURSE DEVELOPMENT

Quantitative Reasoning, MAT1030. Development of core curriculum, including video presentations of each topic, homework assignments, group activities, student projects, and examinations. Integrated technology, such as Excel and OneNote, into the course. Designed course for both face-to-face and online implementations. Course has been used by several faculty members since Spring 2017. **2016**
Quantitative Reasoning – First Year Seminar Version, MAT1030. Development of core curriculum, including print and video presentations of each topic, homework assignments, group activities, discussion boards, student projects, and examinations. Integrated technology, such as Excel and OneNote, into the course. Designed for a hybrid experience – online learning, in-person group work. **2020**
Quantitative Reasoning with Integrated Algebra, MAT1035. Spring 2018. Added necessary components of intermediate algebra to the above course to help students who have a basic understanding of prealgebra succeed in the course. Designed the course for both face-to-face and online implementations. Course has been used by several faculty members since Summer of 2018. **2018**
Quantitative Reasoning with Integrated Algebra – Online course, MAT1035. Development of course, including video and written presentations of each topic, homework assignments, detailed group activities and weekly interactions created, student projects (mostly involved in the group work mentioned above), and examinations. Integrated technology, including Excel and OneNote, into the course. Designed solely for an online environment. Course has been used by several faculty members since Spring 2020. **2020**

PUBLICATIONS

Textbooks

Contributed a new method of factoring to be used as deemed appropriate by the authors Michael Sullivan, III, Kathrine Struve, and Janet Mazzarella for the Publisher (Pearson Education) in connection with the works entitled *Elementary Algebra, Intermediate Algebra, Elementary & Intermediate Algebra, and Developmental Mathematics.*

Published Journal Papers

T. Green, W.H. Goodridge, J.D. Anderson, E. Davishahl, D. Kane, "*Comparing the Effectiveness of Newer Linework on the Mental Cutting Test (MCT) To Investigate Its Delivery in Online Educational Settings*", International Education Studies, vol. 16, no. 4, 2023.
D.J. Lee, J.D. Anderson, and J.K. Archibald, "*Hardware Implementation of Spline-based Genetic Algorithm for Embedded Stereo Vision Sensor Providing Real-time Visual Guidance to the Visually Impaired*", special issue on "Signal Processing for Applications in Healthcare Systems (AHS)" of the EURASIP Journal on Advances in Signal Processing, vol. 2008, doi 10.1155/2008/385827, 10 pages, June 2008.

Refereed Conference Proceedings and Presentations

R. Kamali, J. Anderson, "*Comparison of Virtual Reality: Effects on Student Learning Using Virtual Technology on Nanotechnology Education*," Proceedings of the 2019 ASEE Annual Conference & Exposition, Tampa, Florida, June 18, 2019.
J.D. Anderson, D.J. Lee, and J.K. Archibald, "*Embedded Stereo Vision System Providing Visual Guidance to the Visually Impaired,*" The Third IEEE/NIH Life Science Systems and Application Workshop (LISSA), p. 229-232, Bethesda, MD, USA, November 8-9, 2007.
J.D. Anderson, D.J. Lee, B.B. Edwards, J.K. Archibald, and C.R. Greco, "*Real-time Feature Tracking on an Embedded Vision Sensor for Small Vision-guided Unmanned Vehicles,*" The 7th IEEE International

Symposium on Computational Intelligence in Robotics and Automation (CIRA), p. 55-60, Jacksonville, FL, USA, June 20-23, 2007.
J.D. Anderson, D.J. Lee, and J.K. Archibald, "*Hardware Implementation of Feature Density Distribution Algorithm for Autonomous Robot*," Proceedings of The 31st Annual Conference of the IEEE Industrial Electronics Society (IECON), p. 357-362, Raleigh, NC, USA, November 6-10, 2005.

Conference Proceedings and Presentations
J. Anderson, M. Aeschbacher, "*The Local Impact of Teaching Voting Theory*," 2022 VMATYC Virtual Conference, April 15, 2022.
J. Anderson, M. Aeschbacher, "*The Local Impact of Teaching Voting Theory*," 2022 Innovative Educators Virtual Summit, March 3, 2022.
W. Athens, J. Premo, K. Andrist, J. Anderson, "*Enabling Students Daily Choice of How to Attend Class: Wisdom from Three Faculty Who Taught HyFlex,*" OLC Education Sessio, September 24, 2021.
R. Brinkerhoff, J. Anderson, "*Group Work in Asynchronous Mathematics Course: Maximizing Benefits and Minimizing Obstacles,"* OLC Innovation Studio Design Thinking Challenge, March 15, 2021.
J. Anderson, M. Aeschbacher, "*Online Group Work for Math Students*," Hawkes Webinar Series, August 21, 2020.
J. Anderson, M. Aeschbacher, "*Corequisite Model*," 2021 Innovative Educators Virtual Summit, March 13-14, 2020.
D. Pedrotti, D. Butler, J. Anderson, S. King, "*Building a Bridge to STEM: Transition Programs for Math Readiness*," STEMTech Conference 2015, Phoenix, AZ, November 1-4, 2015.
J. Anderson, *"Flipping the Classroom: Integrating Videos, Online Instruction, and Classroom Lectures,"* SWADE Conference 2014, Orem, UT, October 30-31, 2014.
J. Anderson, *"Flipping the Classroom: Integrating Videos, Online Instruction, and Classroom Lectures,"* UMATYC Conference 2014, Orem, UT, September 20, 2014.
J. Anderson, K. White, "*Lessons Learned from a First Mathematical MOOC*," AMATYC 2013, October 31-November 3, 2013.
J. Anderson, D. Wiberg, "*Synthetic Factoring: Success Focused Technique for Factoring Trinomials,*" SWADE Conference 2012, Salt Lake City, UT, October 31-November 1, 2012.
J. Anderson, "*A Student Centered Approach to Factoring Trinomials*", UMATYC Conference 2012, Ephraim, UT, September 15, 2012.
J. Anderson, D. Wiberg, "*Helping All Students Master Factoring"*, Southern Nevada Math and Science Conference, Las Vegas, NV, January 20-21, 2012.
J. Anderson, D. Wiberg, "*Student Centered Factoring,*" ALADE Conference 2011, November 3-4, Alabama State University, Montgomery, AL.
E. Thompson and J. Anderson, "*Redesigning the Student Learning Environment*," Assessment Best Practices Conference, UVU University College, Orem, UT, USA, March 14, 2011.
J.D. Anderson, D.J. Lee, R.B Schoenberger, and B.J. Tippetts, "*Using Real-time Vision to Control a Convoy of Semi-Autonomous Unmanned Vehicle,"* AUVSI's Unmanned Systems North America, online proceedings, Orlando, FL, USA, August 29-31, 2006.
J.D. Anderson, D.J. Lee, R.B Schoenberger, Z.Y. Wei, and J.K. Archibald, "*Semi-Autonomous Unmanned Ground Vehicle Control System*," SPIE International Symposium on Defense and Security, Intelligent Computing: Theory and Applications III, vol. 6230, 62301M, Orlando, FL, USA, April 17-21, 2006.
C. Archibald, E.S. Millar, J.D. Anderson, J.K. Archibald, and D.J. Lee, "*A Simple Approach to a Vision-guided Unmanned Vehicle,"* SPIE Optics East, Robotics Technologies and Architectures, Intelligent Robots and Computer Vision XVIII, vol. 6006, p. 210-220, Boston, MA, USA, October 23-26, 2005.
J.D. Anderson, D.J. Lee, and J.K. Archibald, "*FPGA Implementation of Vision Algorithms for Small Autonomous Robots,"* SPIE Optics East, Robotics Technologies and Architectures, Intelligent Robots and Computer Vision XVIII, vol. 6006, p. 401-411, Boston, MA, USA, October 23-26, 2005.

MEDIA CONTRIBUTIONS

Learning Intermediate Algebra YouTube site, https://www.youtube.com/channel/UCW11H80p2lJzMXFPhyiAHcg. Contains videos covering every topic found in UVU's Intermediate Algebra (MAT1010) course.
Learning College Algebra YouTube site, https://www.youtube.com/channel/UCnzjrvgsNYn3fISxtid8ynQ. Contains videos covering every topic found in UVU's College Algebra (MATH1050) course along with preparatory material from the preceding course.

PROFESSIONAL SERVICE

Faculty Senate Parliamentarian	**2020-Present**
Academic Program Assessment Committee, Faculty Chair	**2020-2023**
ALEKS Faculty Coordinator	**2019-Present**
Mathematical and Quantitative Reasoning Department Assessment Committee Chair	**2018-Present**
Re-envisioning the Undergraduate Experience Committee (RUEC) Member	**2018-2020**
Faculty Senate Vice President	**2017-2019**
Council of Academic Standards, Chair	**2014-2017**
Academic Technology Steering Committee, Faculty Senate Representative	**2014-2017**
Faculty Senator, University College at Large	**2013-2015**
Vendor Organizer, 2013 South-West Association of Developmental Education Conference	**2013**

PROFESSIONAL DEVELOPMENT ACTIVITIES

Higher Education Authority workshops through the Office of Teaching and Learning at Utah Valley University	**2016-Present**
Utah Valley Senior Executive Leadership Forum (UVSELF), Utah Valley University Office of the President	**2015-2016**
Ethics Across the Curriculum, Utah Valley University Ethics Center	**2014**

SKILLS AND QUALIFICATIONS

Strong oral and written communication in English
Working knowledge of statistical analysis
Working knowledge of Microsoft Office produces, including Excel, Word, and OneNote
Working knowledge of MATLAB and Labview
Working knowledge of C++, C-Sharp, and Java

RESEARCH INTERESTS

Appropriate uses of technology in education, including automated homework, simulations, Excel, OneNote, etc.
Flipped classrooms and its effect on student engagement and learning.
Student engagement in coursework.
Engaging student learning using applications of mathematics in the real world.
Effects of group work on student engagement and learning.

SUPERVISED UNDERGRADUATE COMPETITIONS

Intelligent Ground Vehicle Competition, BYU	**2005-2006**