Using virtual reality as a platform for developing mental models of industrial systems

by

Robert Joseph Slezak

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Human Computer Interaction

Program of Study Committee:
Nir Keren, Major Professor
Stephen Gilbert
Mani Mina

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this thesis. The Graduate College will ensure this thesis is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2020

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank my major professor, Dr. Nir Keren, and my committee members, Stephen Gilbert and Mani Mina, for guiding and supporting me throughout the course of this research, especially in these unprecedented times.

In addition, I would also like to thank Tor Finseth for his advice in graduate college and getting me started on academic writing and to Peter Carlson for assisting me with graphics design, 3D modeling, user interface and bug testing, and for his assistance and feedback on my research. Without their assistance, this research would have not been possible.

In addition, I would also like to thank Dr. Saxon Ryan, Matthew Harvey, and Alex Wiley for assisting me with the analyses.

Finally, I would like to thank my father for supporting me throughout my journey at Iowa State University; his advice and contribution to my academic and daily life have been immeasurable and without him, this thesis would not exist. This serves as a tantamount to my success in graduate college, but also a reminder to always stay by my father's side.

ABSTRACT

To effectively design, build, and interact with industrial systems, engineering and technology students must come prepared with a robust understanding of industrial systems. Developing proper understanding of industrial system is complex and daunting process. To understand industrial systems, students must develop mental models for systems; mental models are dynamic, mental representations of what users know and how they perceive the real world around them. Research on mental models of systems distinguish among four notions of mental models. These notions are *Device Topology*, a notion representing the level of understanding of the structure of a device or a system, usually comprised of individual components; the *Envisioning*, a notion representing the level of understanding of the components' function in the device or system out of context; the *Causal Model, a* notion representing the level of understanding the device's or system's purpose or overall function; and, the *Simulation*, a notion representing the level of understanding how the device or system behaves under specified conditions. In this research, the notion of *Device topology* was divided to two sub-notions to better reflect on the complexity of industrial systems in compassion to devices. These sub notions were titled *Process System topology*, a notion representing components in a system that affect the process, and *Service System topology*, a notion representing components that are servicing the process components.

In recent years, virtual reality (VR) became accessible and attracted interest as a potential learning tool. However, the extent to which VR contributes to positive and enhanced learning remained inconclusive. Multiple studies found that, despite offering higher proper sense of presence, learning outcomes with VR was worse than the learning outcomes with non-immersive technologies; other studies reported that VR provided enhanced learning outcomes in comparison to non-VR instruction. Studies on VR-based learning focused on various topic domains, which seems to hinder reaching generalizable conclusions on the merit of learning in VR.

This thesis pursues two overarching research questions: (1) can interacting, designing, and building systems with interactive VR applications enhance students' mental models of industrial systems; and, (2) Can level of presence predict level of notions of mental models of systems. The two overarching research questions were evaluated using two VR applications titled Cooling Water Virtual Reality (CWVR) and System Designer VR (SDVR). Engineering and technology students participated in experiments with CWVR and SDVR. Students were instructed to explore a prefabricated cooling water system with CWVR and to design and build an industrial system based on task specifications with SDVR. The results demonstrated that students that began with designing and building a system with SDVR and then interacted with the prefabricated CWVR had a modestly higher levels of notions of *Process* and *Service System Topology* and a significantly higher notion of *Causal Model* of the cooling water system in CWVR, in comparison to students that interacted with the prefabricated CWVR without previous experience with SDVR. The results also demonstrated that *presence* was significantly associated with the Service System Topology notion of mental model but not with other notions. There were other significant relationships among interactivity parameters with the various notions; however, a review of these results led to only a causal explanation for the relationships with presence.

The conclusions offer that, potentially, a single engagement with system resulted with a modest shift in the notions of mental models of systems, and that an extensive engagement with an application such as SVDR may results in overall significant elevated levels on all notions.

Further, the lack of significant relationships between presence and notions of mental models, other than with *Service System Topology*, may be explained by the overall number of participants with high-level *presence*.

CHAPTER 1. INTRODUCTION

The primary objective of this research was to determine if specifically, designed VR application can use as a tool for improving students' mental models of industrial systems. A secondary objective was to assess the relationship between level of presence in the specially designed application and the level of mental models of industrial systems. Following a review of relevant literature.

1.1 Motivation

Technology and engineering students are likely to engage with industrial systems in their professional lives. Industrial systems, however, can be a challenging subject to learn. This challenge stems from the complexity of systems. To understand industrial systems, engineering and technology students must understand the following aspects of the systems: the structural and component composition (e.g., the components used in the industrial system), the way components function out of the context of the global system, the interaction between components in the system and their interaction with the environment, and system behavior due to various changes in operations and other conditions. Industrial systems may utilize tens of major process components (e.g., pumps, cooling towers) and thousands of service components (e.g., valves, sensors), making it difficult to understand its operation if each component and its role are not understood. Without understanding the critical aspects above, students may not be prepared to work with industrial systems (e.g., operating, designing, building, and interacting) in real-world applications.

The knowledge or understanding of the fundamental aspects of industrial systems to allow performing industrial system-related tasks (e.g., operation, maintenance) are referred to as *mental models of industrial systems*. Mental models are cognitive, conceptual models of one's self and the world, that is developed through experience, training, and instruction (Norman, 1983; 2013). Mental models act as cognitive guides to predict outcomes, achieve objectives, and improve understanding of the world (Norman, 2013). A person with a robust mental model of its surroundings can predict events at that moment. Mental models speculate events constantly and change or evolve overtime with new knowledge or observations when the actual outcome does not match the speculated outcome. Mental models are unstable, losing information overtime as new knowledge is gain, describing its dynamic nature. Mental models are often imperfect and not technically accurate; however, if the mental model is sufficiently functional or a causal model exists, user can perform actions to complete a task (Norman, 1983).

Traditionally, students in engineering and technology classes, work with industrial systems represented through 2D diagrams, which typically show component composition and mechanical connections. 2D diagrams, however, are not interactive, dynamic, and do not provide a full sense of scale. To further improve knowledge on industrial systems, physical access and interaction with industrial systems is recommended. However, in educational settings, it is difficult to access and likely cost prohibitive to have an industrial system for learning purposes, making it becomes an infeasible option. Without physical access to industrial systems, engineering and technology students lose learning opportunities to visualize, interact, and experience industrial systems. Without those learning opportunities, developing mental models of industrial systems is a major challenge. VR, however, offers an alternative for presenting a full scale, interactive, and three-dimensional industrial system.

VR is a visualization platform that can provide fully scaled, interactive, and three-dimensional objects in the virtual environment and potentially improve visual learning outcomes (Brelsford, 1993; Rupp, et al, 2019). With VR, the user can interact with virtual objects in the environment and view objects at angles difficult to replicate on a desktop or laptop computer (Puschmann et al, 2016). VR has been commonly used in training applications to train trainees without using physical equipment, and thus reducing costs and eliminating risks to life (Brooks, 1999). Recently, VR has become a topic of interest in education, especially with the rise of consumer VR systems; this interest can be attributed to elements such motivation and interest. Parong and Mayer (2018) and Makransky and colleagues (2019) present aspects motivational and interest theories in learning with VR technology. Researchers also report that presence in VR, the sense of being there (Slater et al, 1995), may contribute to positive learning outcomes and assist in completing learning tasks (e.g., Mikropoulos & Natsis, 2011). Studies in multimedia learning, however, found that VR had mixed results on knowledge gain. Several studies found that the knowledge gain in VR was worse than the gain in other medias such as PowerPoint and desktop computer simulation software (Parong & Mayer, 2018; Makransky et al, 2019). One other study found that the knowledge gain was greater than the instructor-led classroom (Webster, 2015). While it seems as the merit of VR as a learning facilitating tool is not clear, the level of merit of learning with VR technology varies by topic domain or is knowledge specific and thus the conclusions so far cannot be generalized. The review of literature of learning with VR above does not provide sufficient evidences to deter attempts for exploring developing mental models for system in VR.

This thesis explores whether performing sophisticated and interactive tasks such as designing and building industrial systems with specially tailored VR applications can enhance developing mental models of industrial systems and whether presence in these applications can predict level of mental models in engineering and technology students.

1.2 Thesis Organization

Chapter 2 reviews the relevant literature. Chapter 3 provides the overarching research questions. The methodology of this research is described in Chapter 4. Chapters 5 and 6 present the results and discussions for the two research questions. Chapter 7 presents a summary and future direction.

CHAPTER 2. LITERATURE REVIEW

This chapter details the background research that is focused on the following topics: 1) VR technologies and applications, 2) mental models and, and 3) presence and its role in VR and studies.

2.1 Virtual Reality

2.1.1 Defining Virtual Reality

Virtual reality (VR) is an experience in which the user is immersed in a virtual world that is either identical to the real world or a fictitious world (Brooks, 1999). In VR, the environment, objects, and elements are completely virtual as explained in the reality-virtuality continuum (Milgram et al, 1994).

VR immerses the user into the virtual environment (VE) using technologies such as head-mounted displays (HMDs) and CAVE™-based surround projection systems and utilize perpetual devices such as game controllers to manipulate the VE and its elements (Brooks, 1999; Havig, 2011). Some researchers and developers referred graphical applications on computers and mobile phones as forms of VR, but those devices do not surround the user with the VE and present the illusion of immersion (Brooks, 1999; Havig, 2011). To distinguish itself from graphical applications on mobile phones and computers with displays, some researchers coined the term *immersive VR* to describe immersive technologies such as HMDs and CAVE™-based systems (Brooker, 1999; Parong & Mayer, 2018; Makransky et al, 2019). Throughout this thesis, VR refers to *immersive VR* to avoid redundancy; other technologies described in this thesis are deem *non-immersive*.

2.1.2 Virtual Reality Technologies

As stated earlier, HMDs and CAVE^{TM}-based systems are commonly used to immerse the user in VR. HMDs are single user devices that are worn on the user's head and the VE is display through two lenses for each eye and changes view when the user's head moves. CAVE^{TM}-based systems are visual surround systems; they utilize walls with projectors or surrounding screens to project the surrounding VE and can hold multiple users (Cruz-Neira et al, 1993). HMDs and CAVE^{TM}-based systems have their own advantages and disadvantages. CAVE^{TM}-based systems offer higher visual resolutions and more lifelike proprioception than HMDs at the expense of high operation and maintenance cost while HMDs are small, portable, and less expensive (Cruz-Neira et al, 1993. Havig et al, 2011). Over the years, HMDs evolved to have higher display resolution, wider field of view, and better ergonomics. Consumer HMDs such as the HTC Vive and the Oculus Rift began to appear in the mainstream market with starting price ranges in the hundreds and can provide higher visual equity and more complex environments with deskop or laptop computers (HTC Corporation, n. d.; Oculus, Inc, n. d.). Mobile-based HMDs such as the Oculus Go also appeared in the mainstream market and do not require a computer for graphical processing, although the graphical processing was limited as it uses hardware akin to smartphone hardware, head tracking was limited to three degrees of freedom, and only supported one controller (Oculus, Inc, n.d.). However, in 2019, the Oculus Quest, a more advanced mobile-based VR HMD, was released and resolve the limitations of mobile-based HMDs; it was designed with higher graphical capabilities and can achieve near computer-powered VR experiences without a computer (Oculus, Inc, n.d.). With the Oculus Quest, the HMD is not tethered to the computer, allowing the freedom to move around. Additionally, two controllers with six degrees of freedom facilitate more natural interaction, increasing the level of immersion and likely to increase the sense of presence. These favorable factors lead to selecting the Oculus Quest as the VR platform for the study herein.

2.1.3 Virtual Reality in Education

In education, motivation is the willingness to engage and learn from lessons or tasks and overcome obstacles that are hindering the understanding of content (Parong & Mayer, 2018). Student motivation in VR is driven by two theoretical concepts: interest theory and self-efficacy theory. Interest theory explains if the learning topic is interesting and perceived as valuable, students will work harder towards to completing lessons and tasks; in other words, interesting and valuable learning content drives motivation in students (Harackiewicz & Knogler, 2017; Parong & Mayer, 2018). Harackiewicz and Knogler (2017) describes learning and attention as feeling ‘effortless’ when content is interesting. Schiefele and colleagues (1992) found that learning outcomes and subject interest are positively correlated in their meta-analysis on interest theory research. Self-efficacy theory posits that students who believe they have the academic horsepower to handle a task at hand will work harder on accomplishing their goal (Schunk & Pajares, 2009; Parong & Mayer, 2018). Students will first judge their self-efficacy (e.g., I can do this, or I can’t do this), which affects their engagement in the task or lesson (e.g., I will work hard to on this task) and determines the students’ motivation and learning outcomes (Schunk & Pajares, 2009; Parong & Mayer, 2018). The use of VR for facilitating motivation in education can be described to make learning content more interesting and valuable and provide immediate feedback for students’ self-efficacy.

Researchers are also interested in utilizing VR for educational use since it offered the sense of presence. Mikropoulos (2006) found that pupils reported a high sense of presence due to interaction with virtual beings and self-representation in a simulated ancient city. In a ten-year review on VR in education, Mikropoulos and Nasis (2011) reported that researchers considered presence to be a necessary component for learning in VR. They also revealed researchers reported that presence contributed to learning outcomes, but the extend on how it contributed remained unknown.

VR is considered by some researchers as a multimedia learning tool as it stimulates multiple sensors (e.g., hearing, vision, etc.) to deliver learning content and instruction to the working memory. Information in working memory is then processed and organized before integrating with verbal and pictorial models and prior knowledge from long-term memory (Mayer & Johnson, 2008; Mayer, 2017). This process is described by the *cognitive theory of multimedia learning*.

Previous research in multimedia learning has focused on developing principles for properly designing multimedia material for education and maximizing its effects on learning outcomes. Examples of multimedia learning principles created for designing multimedia content include the multimedia (learning outcomes are improved when words and pictures are present simultaneously rather than alone), coherence (excluding extraneous material leads to better learning outcomes), and redundancy (concurrent on-screen text, narration, and graphics generate extraneous processing and limit learning outcomes if on-screen text hinders essential processing) principles (Mayer et al, 2001; Mayer & Johnson, 2008; Mayer, 2017). These examples are a snippet of the multimedia learning principles designed through research. Multimedia learning principles were then pursed in various multimedia formats including animation clips, websites, and pedagogical agents to study the impact on learning outcomes (Mayer & Moreno, 1998; Mayer et al, 2001; Moreno and Mayer 1999; Park et al, 2015; Mayer, 2017; Wang et al 2018). Patwardhan and Murthy (2015) explored the impact of adding interactions on learning in visualizations and found that higher levels of interactions did not lead to higher gain in conceptual knowledge over non-interactive visualizations, but visualizations with interactivity features can improve understanding and applying procedural knowledge. However, research found that despite VR offering more presence, motivation, and satisfaction with learning content, participants in VR groups had less knowledge gain and learning performance compared to non-immersive technologies and physical science labs (Parong and Mayer, 2018; Makransky et al, 2019). Parong and Mayer (2018) showed that PowerPoint version of the VR application resulted in higher knowledge gain than the gain from using the VR application directly, pointing to the coherence principle violation. However, in a follow up experiment, Parong and Mayer (2018) demonstrated that adding generative learning (e.g., summarizing lessons) to VR improves knowledge gain. Makransky et al (2019) revealed that VR users suffered from heavier cognitive workloads and performed worse than desktop users.

The results from multimedia learning research demonstrated that there is no consensus regarding the benefits of using VR in learning and other education endeavors. Similarly, multimedia-based learning may not be generalizable to other subject-domains, lessons, or other technologies. For example, the results of teaching the human body (Parong & Mayer, 2018) and simulating a science lab in VR (Makransky et al, 2019) may not apply for teaching industrial systems or other subject matters in the engineering domain. Contrary to the consensus, Alhalabi (2016) demonstrated that three VR groups had higher exam scores than the No-VR group in engineering education, demonstrating VR potential in that domain. Dinis et al (2017) found that technologies like VR help generate interest in civil engineering where students perceived VR to be important and often essential to understanding concepts and motivation for learning. Given that industrial systems are part of engineering domain, it is worthwhile to explore if VR can enhance students’ understanding of industrial systems.

2.1.4 Applications of Virtual Reality

VR was utilized to simulate real world scenarios, places or objects. Other than educational use, VR was also used in training. Training in VR became an alternative to traditional, physical solutions since it was more cost-effective to simulate real world scenarios and train within the simulation. Brooks (1999) experienced a 747 simulator from British Airways, which consisted of a pilot cockpit with controls and display and can be re-used to train trainees, eliminating the need for multiple 747 planes for pilot training. Wyk and Villers (2009) performed a study on training miners in Africa in VR to reduce injuries and fatalities and found it was more cost effective than real-life training, eliminates the risk of injury or death for trainees, and trainees found the simulation engaging. Ochs et al (2019) developed a conversional agent that simulates the patient on desktop , HMD, and CAVE™-based systems and measured presence to validate the agent's capability of training doctors' social skills to break bad news; they found that HMD and CAVE™-based systems reported higher presence than desktop systems and that experts were more engaged than novices.

VR was also used to create designs or conduct design reviews. Brooks (1999) explained that Daimler-Chrysler and John Deere used VR to improve ergonomics and create mockup designs respectively, saving money in prototype costs. In risk assessments on machine tools, Puschmann et al (2016) found that compared to document-based and CAD-based risk assessments, VR-based risk assessments helped the participants discover more risks and design flaws.

Simulating real world scenarios or objects without needing physical objects or endangering the student's life is an appealing characteristic in training and educational applications. Trainers and educators can utilize VR to simulate scenarios where physical access is not feasible, cost prohibitive, or even dangerous. For industrial systems, it will be difficult for students to gain physical access to them. A VR simulation of industrial systems is cost-effective and accessible, unlike physical industrial systems.

2.2 Mental Models

2.2.1 Overview

As stated earlier, mental models are defined as mental representations of the real-world served as cognitive guides to predict outcomes, achieve goals, complete tasks, and improve knowledge (Norman, 1983; 2013). Depending on the domain of discussion, the subject domain, and the researchers' objective, mental model definitions may vary to some degree. Since mental models can encompass many subject domains and domains of discussion, this overview will focus on examining changes in mental models and mental model measurement methodology.

Some researchers examined cognitive changes and structural differences through experience gain in mental models or studied if mental models can predict user behavior. Hegarty et al (2013) examined cognitive changes in students' mental models and problem-solving strategies in organic chemistry through strategy training; they found that users strategized multiple ways to solve problems after training. Furlough and Gillan (2018) explored structural differences in mental models in three levels of experience, novice, medium, and expert, but failed to find significant differences in the density of mental models among various experience levels. However, Furlough and Gillan (2018) identified that novice players had more natural language connections than medium and expert players, while expert players had more procedural connections, subnetworks, and abstract links than medium and novice players and proposed a three-stage theory on mental model construction to define level of experience. Kang et al (2015) explored whether mental models of the Internet affect users' decisions regarding privacy and security and failed to find significant connections between technical knowledge and privacy and security decisions.

Earlier research on mental models show that variety of measurement instruments and techniques were utilized to gauge strength of mental models. Common tools and methodology used for measuring mental models include knowledge tests, think-aloud troubleshooting protocols, laddering interviews, relatedness ratings, and diagramming. (Rowe & Cooke, 1995; Holman, 2011; Revell & Stanton, 2014; Chen et al, 2015; Jalmo & Suwandi, 2018). According to Rowe and Cooke (1995), there was no consensus among researchers on a standard for mental model measurements; the authors claimed that measurement techniques and instruments measure different areas of mental models. Rowe and Cooke (1995) found that laddering interviews, relatedness ratings, and diagramming could predict measuring mental model performance for troubleshooting, over think-aloud troubleshooting protocols. Rowe & Cooke (1995) also discovered that laddering interviews and relatedness ratings are independent as they focus on different areas of mental models. Some studies had student participants draw diagrams of search engine strategies and home heating systems to demonstrate their understanding of these systems (Holman, 2011; Revell & Stanton, 2014). Chen et al (2015) utilized eye tracking to measure performance differences between a static 3D representation and a dynamic 3D representation on students. Jalmo & Suwandi (2018) had students perform writing and drawing tests on genetic concepts to find which test offers higher levels of mental models.

Researchers collected and organized ground and scholarly information from expert sources and common user information and beliefs from user interviews to form expert and user mental models respectively (Cox et al, 2003; Austin et al, 2020). In this domain, researchers cross-examined expert and user mental models to find key differences in understanding among various topics, which generally indicated misunderstood information or weaknesses in understanding (Cox et al, 2003; Austin et al, 2020). Lau and Yuen (2010) compared mental models of sorting algorithms from males and females, and concrete and abstract learners with expert mental models; they found that mental models in females and concrete learners were closer to expert understanding. Cox et al (2003) found misunderstandings of chemical safety from cross-examination of expert and user mental models and created communication content for the chemical workers to encourage safety. Austin et al (2020) discovered no significant insight on reducing risks of electrocution on expert-user mental model examination but found it can identify ignored safety practices.

In academic settings, Hestenes and colleagues (1992) created the force concept inventory from pre-tests to identify weaknesses in conceptual understanding on Newtonian mechanics and performed post-tests to determine if the concept inventory improved students' conceptual models. Jalmo and Suwandi (2018) found students possess a higher level of mental model of genetic concepts through open response writing tests over drawing tests, concluding it may be related to low spatial ability. Holman (2011) finds that millennial students lack a conceptual model of search engine mechanisms that prevent them expanding into deeper search methods in academic research and those with more robust mental models perform more complex searches.

2.2.2 Notions of Mental Models of Systems

Norman (2013) describes that users form mental models of systems by forming conceptual models or simplified representations of how systems work and then combine those models with previous knowledge, information around the environment, and information on the system to form a system image, which determines how effectively the user can interact with the system. Since our interest in mental models involves the engineering domain, we would want students to form robust models of mechanical systems. In this scenario, the mechanistic mental model is the most appropriate mental model domain to investigate.

The motivation for researching mechanistic mental models was to understand the level of veracity between the cognitive representations and the physical mechanical devices or systems (De Kleer & Brown, 1983). According to De Kleer and Brown (1983), mechanistic mental models are comprised of four notions:

1. *Device Topology* represents the level of understanding of the structure of a device or a system, usually comprised of individual components.
2. *Envisioning* represents the level of understanding of the components' function, not in the context of the whole system, but in a localized fashion.
3. *Causal Model* represents the level of understanding the device's or system's purpose or overall function.
4. *Simulation* represents the level of understanding how the device or system behaves to a level where problem solving can be applied.

Revell and Stanton (2014) provided support for the four notions above. They explored the *Device Topology* and *Causal Model* notions to examine mental models of a UK home heating systems among six participants and demonstrated they could distinguish among three drastically different mental models among their subjects from the technical model and could predict energy consumption behavior from the participants' mental models.

This study explores mental models of industrial systems following the notions of mechanistic mental models. De Kleer's and Brown's (1983) definition of mechanistic mental models and the model's notions are adopted with minor adjustments. In this study, the structure of the industrial systems will be referred as *System Topology* instead of Device Topology. This study will also examine two sub-notions of the System Topology: *Process System Topology* notion and *Service System Topology* notion. *Process System Topology* refers to the industrial components that change or facilitate the process for which determines the industrial system's function. For example, a power plant consists of a mechanistic power generation unit that mechanically generates energy, a generator that converts the mechanical energy to electrical energy, and heat exchangers or heaters to ensure that the industrial system is at the desired operational temperature; these generators, heat exchangers, and heaters are examples of process components since they determine the power plant's function. *Service System Topology* refers to industrial components that are put in place to service the process components of the industrial system. Continuing the power plant example, the power plant may include service components such as cooling pumps, valves, and instruments like temperature sensors so that the system can notify other components to change states or so that the individual may perform maintenance. An individual with expertise in power plants will be able to anticipate and then recognize process components of power plant even if the individual has not been in the specific plant yet. Thus, the experienced individual may have a casual representation of systems without seeing them. Recognizing items in Service System Topology require a deeper familiarity with the system. The *Envisioning*, *Causal Model*, and *Simulation* notions are adopted directly from De Kleer's and Brown's (1983) definition.

2.3 Presence

2.3.1 Definition

The general concept of presence, of being physically located in an environment, became dominant in the extended reality arena. Slater and his colleagues (1995) describe presence as a psychological sense of 'being there' in the VE. They also present that presence is highly dependent on the extent to which proprioception match expectation associated with multisensory input and is independent of the match between self-representation and the multisensory input since the user's personal self-model may be contradicted, no matter how real the virtual body is.

Increasing presence, however, can increase self-representation and immersion. Slater et al (1995) posit that presence is a composite of three components: the sense of ‘being there’, the extent to which the VE was the dominant reality, and the extent to which the experience in the VE was similar to visiting a place rather than seeing images of the place. Witmer and Singer (WS) (1998) defines presence as a “subjective experience of being in one place or environment, even when one is physically situated in another” and as “experiencing the computer-generated environment rather than the actual physical locale.” Presence is then treated as a variable defining some quality aspects of VEs. Mikropoulos (2006) also defines presence as the sense of being there but adds that presence is “non-mediation” and an estimation between objective and subjective reality.

As described by the researchers, presence has been widely acknowledged as a psychological sense of “being there” (Schubert et al, 2011). Like Witmer and Singer, other researchers claim that presence is influenced by multiple factors. Heeter (1992) explains that presence encompasses three dimensions: personal, social, and environmental presence. Personal presence refers to the extent to which one feels as if in the VE. Social presence refers to the extent to which the user accepts that other beings (real or virtual) also exist in the world and appear to react to the user. Environmental presence refers to the extent to which the environment itself appears to recognize the user in it and react to the user. According to Heeter, each presence dimension has its own influential factors. Personal presence is influenced by immersion, familiarity, and self-representation. Other users and virtual beings contribute to the social construction of reality (mainly studied in social presence). Levels of environmental presence are dependent on level of user perception that VE is responding to the presence of the user and other beings. Witmer and Singer (1998) states that the following factors that contribute to the sense of presence: control (e.g., degree of control, anticipation of events), sensory (e.g., environmental richness, multimodal presentation), distraction (e.g, isolation, selective attention), and realism (e.g., consistency with real world, meaningful experience). They also posit that presence is both, a characteristic of VEs and a function of individual differences. Schubert and colleagues (2001) state that presence does not have a one-to-one relationship with immersion since there are cognitive factors leading to presence from stimuli perception. They argued that mental model construction and attention allocation are two cognitive processes involved in presence, and thus presence does not directly represent cognitive processes.

Several researchers demonstrated that the sense of presence is defined by the actions permitted in the VE rather than the graphical fidelity of the VE and its structure. Zahorik and Jenison (1998) claimed that accepted definitions of presence above are ill-defined, subjective, and based much upon the VE formalization. Instead, the authors defined presence as a “tantamount to successfully supported action in the environment.” Zahorik and Jenison based their definition on the view of “being-in-the-world” or the fundamental nature of existence and that the real and virtual environments offer affordances or perceived information that informs the users of the actions they can perform. The authors believe that affordances facilitate in the understanding of the environment and its objects instead of mental representation.

2.3.2 Measuring presence

The two most commonly used instruments for measuring presence are the Slater, Usoh, and Steed (SUS) (2000) Questionnaires and Witmer and Singer (1998) (WS) Questionnaires. The first edition of SUS Questionnaires started with three questions focusing on the following areas: the sense of ‘being there’, the extent to which the VE was the dominant reality, and the extent to which the experience in the VE resembled visiting a place rather than seeing images of the place (Slater et al, 1995). The questionnaire was ranking based on 7-point Likert scale. Three more questions were later added to the SUS Questionnaires to allow more refined distinguishing between reality and Virtual Reality (VR) environments (Usoh et al, 2000). Witmer and Signer (1998) devised two sets of questionnaires: Presence Questionnaires (PQ) for measuring presence and Immersive Tendencies Questionnaires (ITQ) for measuring individual ability to get immersed. Slater (1999) was critical of the four factors in WS' instrument suggesting the differences in responses are more likely due to individual differences (e.g., experience, dexterity) and that difference in perception of immersion are not necessarily indicative of levels of presence. According to Witmer and Singer (1998), presence in VEs is a function of individual and VE characteristics, which Slater claims it is impossible to separate between two factors. Slater was also critical of the relationship between presence and task performance in WS, claiming the metrics depend on user interfaces, personal skills, and experience. However, Slater does indicate ITQ is proper for measuring psychological characteristics. Schubert et al (2001) provided evidences supporting Slater's criticism of WS PQ. Factor analyses from two experiments led to a distinction among presence, immersion, and interaction providing that items in WS PQ measure subjective evaluation of the contributing factors rather than presence experiences. Nystad and Sebook (2004) also evaluated SUS and WS presence questionnaires to determine which is superior in measuring presence. Their results showed that SUS questions were positively correlated with personal factors and negatively correlated with performance (more errors, less presence) and WS questions were positively correlated with usability. They found no significant relationship between levels of immersion and presence. Nystad and Sebook (2004) concluded that SUS is more consistent with the concept of presence and WS seems to be more related to technology and interaction factors.

2.3.3 Drawbacks of Presence instruments

Break-in-presence: Slater, Steed, Schwind and colleagues pointed out that users experience a break-in-presence when they transition from the virtual experience to reality and that the re-adjustment to the real world may affect the response to post-experimental assessment, leading to reduced reliability (Slater & Steed, 2000; Schwind et al, 2019). Slater and Steed (2000) then examined an in-situ assessment of presence based on counting the number of times users transitioned from a sense of “presence in the VE” to “presence in the real world.” Their results showed clear positive relationship between their in-situ assessment approach and post-experimental SUS instrument. Schwind et al (2019) examined using presence questionnaires during the VR experience to determine if it reports significantly different presence scores from post-experience assessment. Their results show that there was a significant variance increase in questionnaires when users answered the questionnaires after experiencing an abstract scene and a decrease after experiencing a realistic scene. It also revealed that virtual questionnaires had a negligible effect on users’ workload, pointing that from workload perspective, presence questionnaires are a reliable approach for a VE in-situ assessment.

Distinguishing among environmental platform: Usoh et al (2002) posit that reality offers the ultimate presence and that presence assessment instruments should be able to distinguish presence between reality and a VE that mocks the real world. They utilized SUS and WS in an experiment to measure presence in both real and VEs and found that SUS was marginally significant in distinguishing between reality and its digital representation in VE, while WS found no significant difference at all. They concluded that presence questionnaires may be useful for measuring presence within environmental platform (e.g., real, VE) but is doubtful for comparing across environmental platforms.

Is presence truly experienced: Slater (2004) claims reliance on post-experimental assessment of presence with questionnaires may call the user to form the sense of presence just by asking about it and that questionnaires cannot provide evidences that presence actually existed during experiences in VEs.

Despite the drawbacks above, presence questionnaires are still the decisive tools for measuring presence. For example, Slater (1998) explained that while he does not appreciate presence questionnaires, he uses a questionnaire because it is one of the few measurements out there for presence. Thus, until more observable methodology is available for presence assessment, presence questionnaires remain the choice. Due to the significant criticism of WS' reliance of user subjective assessment and the potential for contaminations due to individual differences, SUS was selected for assessing presence in this research.

CHAPTER 3. OVERARCHING RESEARCH QUESTIONS

The overarching research questions in this study are as follows:

1. *Can VR enhance students' mental models of industrial systems?*
2. *Does presence contribute to the enhancement of students' mental models of industrial systems?*

The following section describes the methodology used in this research. To answer the overarching research questions, a VR framework composed of two VR applications was developed. The first VR application titled Cooling Water Virtual Reality System (CWVR) simulated a prefabricated cooling water system; students could explore the system and interact with the system components. The second VR application titled System Designer VR (SDVR), allowed students to design and build industrial systems. Students were instructed to design and build the industrial system per problem specification with SVDR. The VR applications and the methodology used in this study are presented in detail in the next chapter.

CHAPTER 4. METHODOLOGY

4.1 VR System and Application Development

Due to its autonomous feature and its ability to use six degrees of freedom controls, the Oculus Quest (https://www.oculus.com/quest/) was the VR platform of choice for this research. The CWVR and SDVR were developed as two separate applications. The applications were created with Unity, a cross-platform game engine that is widely used to develop games and 3D applications and widely supports consumer VR Head Mounted Devices. Additionally, two Unity add-ons, Oculus Integration and Virtual Reality Toolkit (VRTK) were utilized to further develop and optimize the VR applications for the Oculus Quest. Figure 4.1 presents the architecture of the CWVR and SDVR.

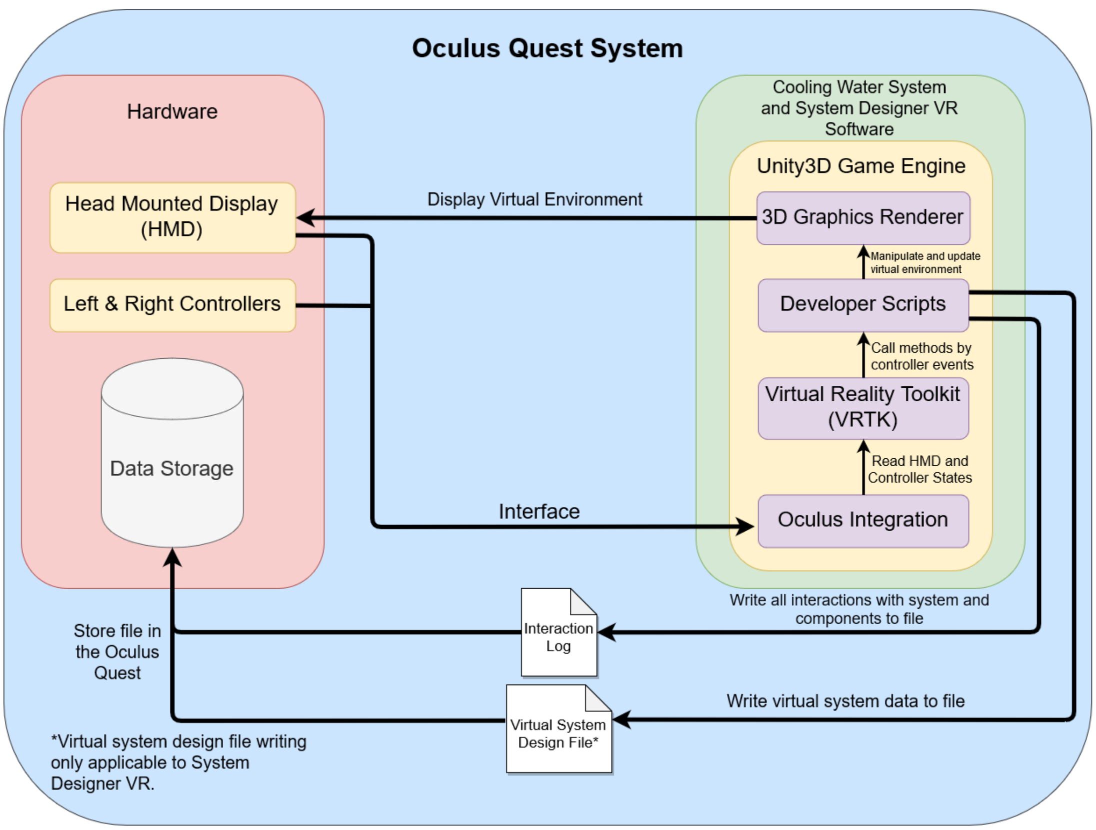

Figure 4.1. Application Architecture of the Cooling Water System and System Designer VR applications.

4.2 Cooling Water System

Figure 2 shows an overview of CWVR. The cooling water used is build according to a general layout of a an industrial cooling system; process components in the cooling water were a reservoir, pumps, a reactor, a cooling tower, and a cooling tower bypass system; service components included various valves, pipes and connectors. Subjects were placed in the CWVR open scene and are told that the objective of their task is to understand how the cooling water system is structured and determine how it functions. Subjects could move around in the scene using the teleporting function with the Oculus Quest controller. They could examine components of the system by grabbing them with the controller and moving them around for visual inspection (see Figure 4.3). When components are released, they slowly returned to their original location in the system. To assist subjects with interpreting the system, 2D images containing a list of industrial components with their schematic presentation, VR image for the components, and descriptions are placed in the scene, behind the system as shown in Figure 4.2. Subjects are told that they are not limited in time. Their task ends when the they notify the research representative; they are completed their task.

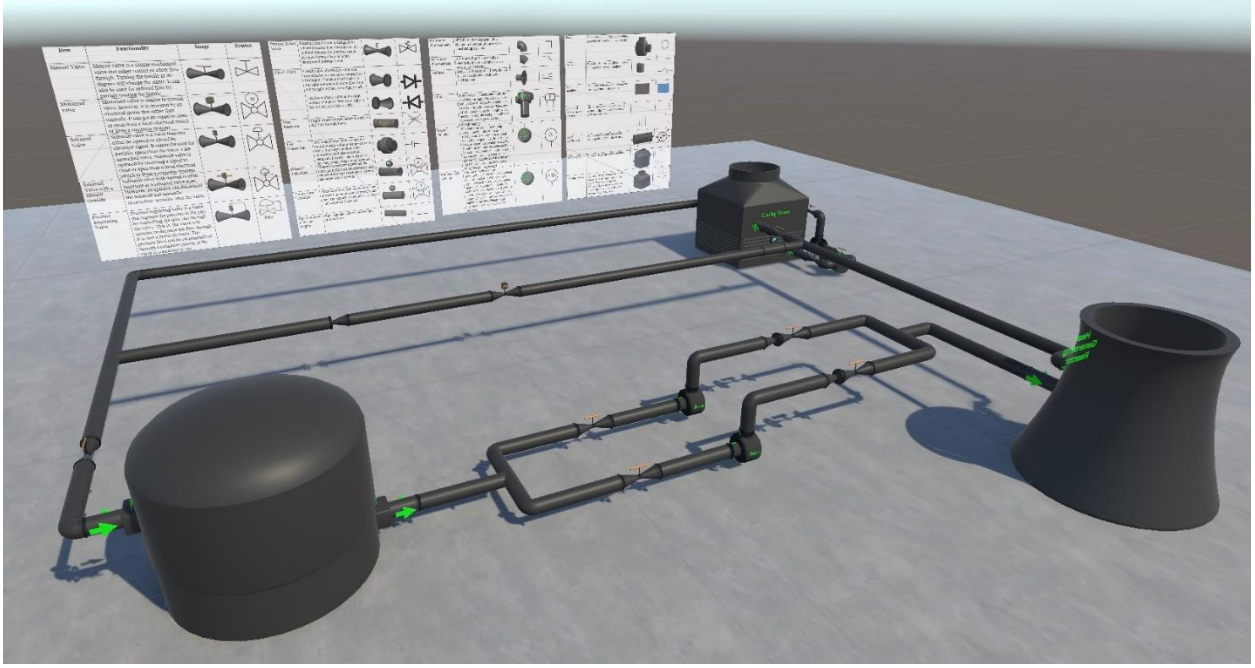

*Figure 4.2***.** Cooling Water Virtual Reality System.

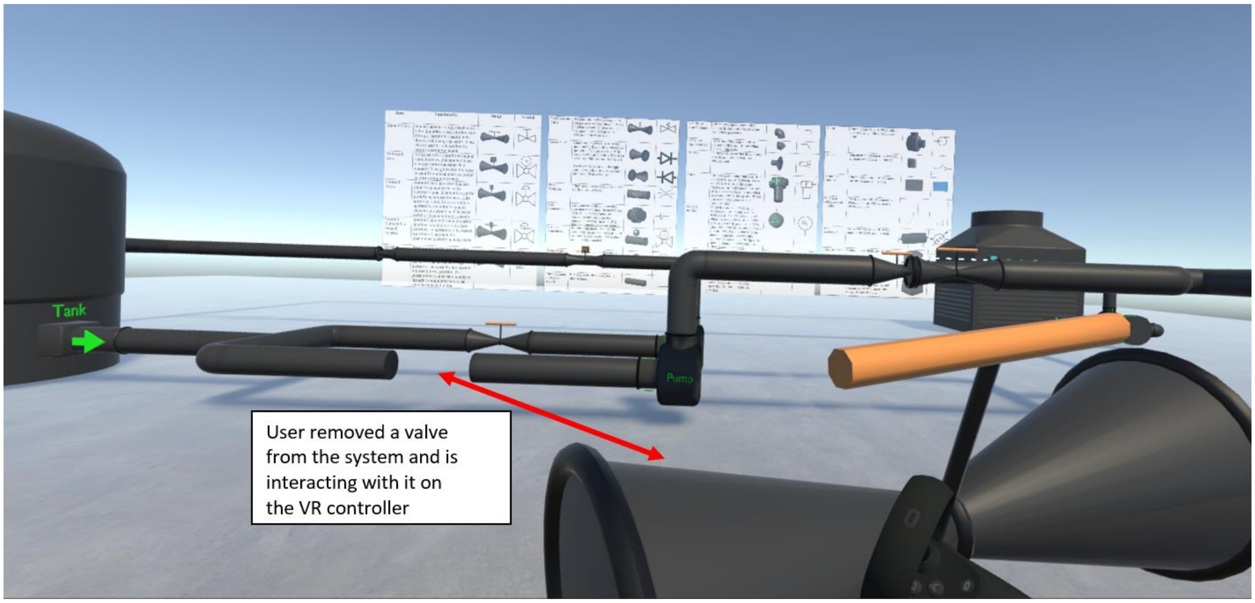

*Figure 4.3***.** User interact with a component in CWVR.

4.3 System Designer VR

The SDVR scene simulates an empty machine room with a steam boiler (Figure 4.4) and two water ports on the wall, one for providing hot water services to customers outside of the machine room, and one for returning water from the customers (Figure 4.5).

Figure 4.4. Steam boiler in the machine room.

Figure 4.5. Water port on the machine room wall.

Unlike the CWVR, subjects in SDVR are tasked to design and build an industrial system with the steam boiler. To build the system, subjects use an interactive component menu to spawn industrial components. Figure 4.6 presents a screen shot of a user interacting with the component menu. Then subjects can connect industrial components together and/or remove them when unneeded. The interactive component menu can be recalled anytime. Subjects could teleport throughout the machine room and grab and interact with system components.

Similarly, to CWVR, 2D images with a list of industrial components and their images and descriptions are posted on the walls. Additionally, assignment instructions are also posted on the wall. Subjects are not limited in time; their task ends when they notify the researchers. Video demonstrating SDVR can be viewed here: https://youtu.be/hErpeptKm9Y.

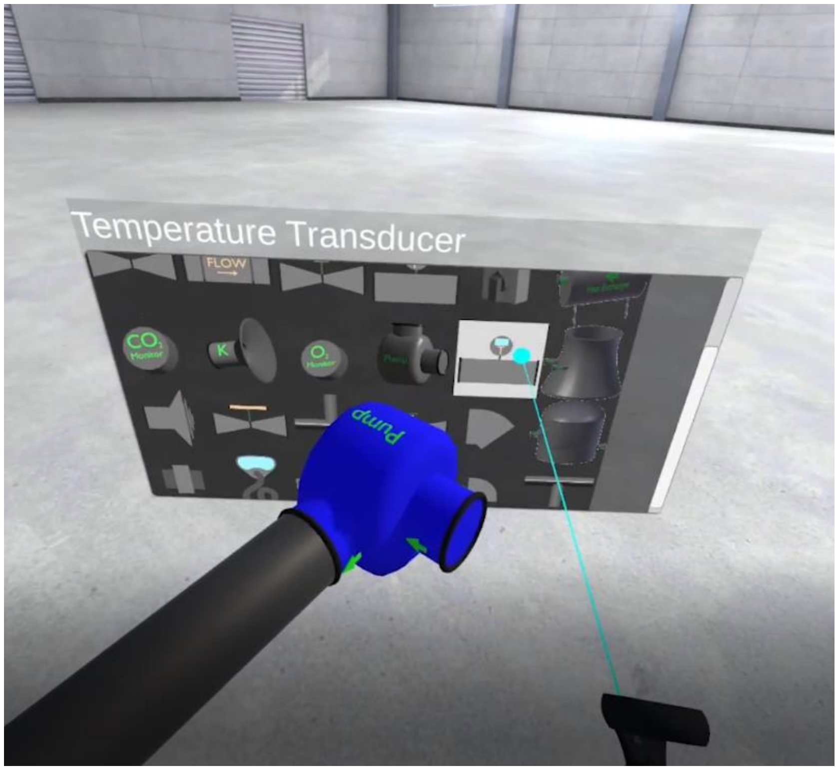

Figure 4.6. Subject interacts with component menu in the scene.

Instructions for the task with SDVR were as follows:

> ***Problem statement:*** *A machine room is built for a fabricating a system that delivers hot water to consumers that require hot water for kitchen services and similar. The machine room is equipped with a boiler that heats the water to a temperature of 300 [ºF] and a pressure of 80 [PSI]. The temperature of the water that go to the consumer should be 180 [ºF] or less. Colder water that return from consumers need to be collected, cooled further, and sent back to the boiler to be reheated again and continue the process.*
>
> *The boiler controls the temperature and the pressure based on information provided from measurements along the pipes. Hot water to consumer is used for food processing and water purity is critical.*
>
> ***Statement of work:*** *You are asked to design and build a system that will facilitate the process above. You will be using the System Designer VR application to build the system as you have done in the VR tutorial. Take screen shots and videos to document your work.*

4.4 Subjects

Subjects are 33 undergraduate and graduate students from various departments at the College of Engineering at Iowa State University. They were recruited from various courses. For participating in the experiment, subjects received 3 extra points to their course grade. Prior to participating in the experiments, subjects reviewed a pre-recorded presentation that introduced them to industrial systems and reviewed the various system components in detail. Subjects could ask any questions if needed. Prior to engaging with the VR experiments, subjects participated in an interactive VR tutorial session that demonstrated to them all functions in the VR scenes from teleporting to interacting with components and spawning components with the interactive component menu. Subjects have been divided to two groups: 17 subjects started with CWVR and then moved to SDVR (Pre-group); 16 started with SDVR and then moved to CWVR (Post-group). The study was approved by the Iowa State University’s Institutional Review Board. One participant withdrew their consent and their data has been excluded from the experimental results.

4.5 Mental Model Assessment

All participants completed a mental model assessment after each task. The mental model assessment instrument consisted of six, essay questions, where subjects are asked to reflect on various aspects of the cooling water in the CWVR and on the system they designed and built in SDVR. These questions were modeled after the notions of mechanistic mental models (De Kleer & Brown, 1983) with the adjustments to systems research as described earlier. Table 4.1 presents the mental model assessment instrument with respect to the relevant mental model notions.

Table 4.1

Mental Model Assessment Questionnaires for Cooling Water System

Notion of Mechanistic Mental Models	Questions
System Topology (includes both Process and Service Topologies)	*What components were used to construct the cooling water system* [for CWVR] */ your system in the machine room* [for SDVR]*? For each listed component, how many were used?*
Envisioning	*What is the role of each component for the cooling water system* [for CWVR]*/ for your system* [for SDVR]*?*
Causal Model	[for CWVR]: *How does the cooling water system keep the water tank from overheating?* [for SDVR]: *How does your system deliver hot water to the consumers and re-use the cool water from consumers?*
Simulation for CWVR (three questions for this notion)	*If the temperature controller reads that the water tank is too high, how will system mitigate the heat in the tank?* *If all check valves fail, what will happen to the cooling water system?* *Replaced the motorized valve with a normal valve. What will happen to the cooling water system?*
Simulation for SDVR (three questions for this notion)	*If the temperature of the hot water that goes to the consumer exceeds 180 degrees F, how will your system handle the excess temperature?* *If the pressure of the hot water exceeds 80 PSI, how will your system handle the excess pressure?* *Replace the manual valves with pipes. What will happen to your system?*

For questions involving the Process System Topology, Service System Topology, and Envisioning notions, responses are rated from 0 to 100; ratings for these notions were based on the number of industrial components the subject remembered correctly (for Process System Topology and Service System Topology) and the number of components' function correctly described for Envisioning.

Subjects' responses for the Causal Model and Simulation notions were rated with either 'Poor', 'Fair', 'Good', or 'Excellent' based on a rating rubric (Appendix A). The ratings received numerical values as follows: Poor=1, Fair=2, Good=3, and, Excellent=4. A review team consisted of two faculty from the Department of Agricultural and Biosystems Engineering, a mechanical engineer from department of Environmental, Health, and Safety, and a graduate student who pursued his M.S. in Human Computer Interaction on this project (also holds an undergraduate degree in Computer Engineering) rated the responses on the notions. All team members are experienced with industrial system safety. Before the assessment began, the members of the team participated in a 90-minute review session on the purpose of the study and all procedures associated with the experiments and the assessment. The members of the team then reviewed the data individually and sent their rankings to the research team. Means from the team ratings were used to analyze the results.

4.6 SUS *Presence* Instrument

Slater et al (1995) defined presence as the sense of ‘being there.’ They posit that presence is dependent on the extent of the match between proprioception and sensory data. Osoh et al (2000) later stated that the ultimate test of virtual environments (VE) is their ability to convey real word experiences, and thus. presence measuring instruments should be able to discriminate between experiences in VE from experiences in the of real world. Usoh et al (2000) then adapted the Slater’s and colleagues’ presence questionnaire (1995) to capture the extent to which presence is comparable to real world experiences by adding three questions. The extended questionnaire is referred to as SUS. SUS maintained ‘loyalty’ to the three aspects of the earlier Slater’s and colleagues’ presence questionnaire. These three aspects can be described as follows:

- **Being there**: The sense of ‘being there’ in the virtual environment
- **Dominance of VE vs dominance reality**: The extent to which the virtual environment becomes the dominant ‘reality’ over the real world
- **VR experience as a place or an image/multimedia**: The extent to which the virtual environment experience is remembered as a place visited in the real world rather than just seen as images or other multimedia format.

SUS Questionnaires were assessed only after the SDVR experience. Table 4.2 presents the modified SUS to refer to the machine room in SDVR. A third column captures the relevant aspect of presence.

Table 4.2

SUS Questionnaires for SDVR

#	Question	*Presence* aspect
1	I had a sense of 'being there' in the machine room. 1 (Not at all) … 7 (Very much)	Being there
2	There were times during the experience when the machine room was the reality for me... 1 (At no time) … 7 (Almost all the time)	Dominance of VE vs dominance of reality
3	The machine room seems to me to be more like… 1 (Images that I saw) … 7 (Somewhere that I visited)	VR experience as a place or an image/multimedia
4	I had a stronger sense of… 1 (Being elsewhere) … 7 (Being in the machine room)	Being there
5	I think of the VR machine room as a place in a way similar to other places that I’ve been today... 1 (Not at all) … 7 (Very much so)	VR experience as a place or an image/multimedia
6	During the experience I often thought that I was really standing in the machine room. 1 (Not very often) … 7 (Very much so)	Dominance of VE vs dominance of reality

4.7 Data Collection and Analysis

SVDR and CWVR included an event recoding algorithm that generated interaction logs. Interaction logs consisted of information on various times, component grabbing, component snapping, and component deletion. All logged events are timestamped to determine the interaction time of the system designing process and the overall simulation time. Additionally, when the students notify the researchers of their completed system in SDVR, the researchers take over the Oculus Quest to save their system designs for further assessment. Student responses to the Mental Model Assessment and the SUS Questionnaires are administered and collected through Qualtrics. All data analysis was conducted using JMP Pro 15 software (SAS Institute Inc., n. d.).

CHAPTER 5. CAN VIRTUAL REALITY ENHANCE STUDENTS' MENTAL MODELS OF INDUSTRIAL SYSTEMS

5.1 Research Question and Hypotheses

The overarching research question is whether a tailored VR framework can serve as a platform for enhancing mental models of industrial systems in engineering and technology students. The overarching research question was broken to five hypotheses, each addressing one of the notions of mental models. Hypotheses and their alternatives are listed below:

Notion of *Process System Topology*:

H_0^1: Notion of *Process System Topology* is not significantly higher in the *Post-group* than in the *Pre-group.*

H_a^1: Notion of *Process System Topology* is significantly higher in the *Post-group* than in the *Pre-group.*

Notion of *Service System Topology*:

H_0^2: Notion of *Service System Topology* is not significantly higher in the *Post-group* than in the *Pre-group.*

H_a^2: Notion of *Service System Topology* is significantly higher in the *Post-group* than in the *Pre-group.*

Notion of *Envisioning*:

H_0^3: Notion of *Envisioning* is not significantly higher in the *Post-group* than in the *Pre-group.*

H_a^3: Notion of *Envisioning* is significantly higher in the *Post-group* than in the *Pre-group.*

Notion of *Causal Model*:

H_0^4: Notion of *Causal Model* is not be significantly higher in the *Post-group* than in the *Pre-group.*

H_a^4: Notion of *Causal Model* is significantly higher in the *Post-group* than in the *Pre-group.*

Notion of *Simulation*:

H_0^5: Notion of *Simulation* is not be significantly higher in the *Post-group* than in the *Pre-group.*

H_a^5: Notion of *Simulation* is statistically higher in the *Post-group* than in the *Pre-group*.

5.2 Description of Dependent Variables

The dependent variables were the ratings on the notions of Process System Topology, Service System Topology, Envisioning, Causal Model, and Simulation. All dependent variables are continuous. The independent variable was the order of exposure to the CWVR, either Pre-group or Post-group.

5.3 Evaluation Results

The following section presents statistical evaluation of each mental model notion by group. Significant difference criterion was set to $\alpha = 0.05$; moderate significance was set to $\alpha = 0.1$. When significance or weak was detected, Cohen's *d* was calculated for effect size. Small, moderate, and large effects are corresponding with *d* values of 0.2, 0.5, and 0.8, respectively (Cohen, 1998). The description provided by a few of the students on certain notions was extremely difficult to interpret; members of the rating team flagged these notions for these students and values for these notions were excluded from the analyses for the corresponding notions.

Table 5.1 presents means and standard deviations for *Process System Topology* by group.

Table 5.1

Process System Topology Notion for Cooling Water System

	n	M	SD
Pre	16	52.27	22.878
Post	16	63.07	19.204

Process System Topology data was examined for outliers, normal distribution, and homogeneity of variance. Dataset met all assumption. Results from a pooled t-test indicate that level of Process System Topology notion in the post-group is moderately significantly higher than this notion in the pre-group, t=1.44566, *p*=.0793. Thus, hypothesis H_0^2 was rejected and the alternative hypothesis H_a^2 is accepted. Effect size for *Process System Topology Notion* was moderate, *d*= 0.51.

Table 5.2 presents means and standard deviation for *Service System Topology* by group.

Table 5.2

Service System Topology Notion for Cooling Water System.

	n	M	SD
Pre	15	54.00	21.974
Post	13	68.45	27.339

Data inspection revealed that neither requirements for outliers nor normal distribution were met. Therefore, Wilcoxon rank sums test was applied on the dataset. Test results indicate that level of Services System Topology notion in the post-group is moderately significantly higher than this notion in the pre-group, z=1.38974, *p*=.0690. Thus, hypothesis H_0^2 was rejected and the alternative hypothesis H_a^2 is accepted. Effect size for *Service System Topology Notion* was moderate, *d*= 0.58.

Table 5.3 presents means and standard deviation for *Envisioning* by group.

Table 5.3
Envisioning Notion for Cooling Water System.

	n	M	SD
Pre	9	54.32	26.32
Post	17	49.02	29.41

Data was examined for outliers, normal distribution, and homogeneity of variance; all requirements met. Results from a pooled t-test indicated no significant main effect, *t*=-0.452519, *p*= .6725. Thus, hypothesis H_0^3 is accepted.

Table 5.4 presents means and standard deviation for *Causal Model* by group.

Table 5.4
Causal Model Notion for Cooling Water System.

	n	M	SD
Pre	16	1.859	0.563
Post	17	2.515	0.886

Inspection revealed data failed to meet outliers and normal distribution requirements. Thus, Wilcoxon rank sums test was applied. Results indicated that *Causal Model* notion is significantly higher in the *post-group* in the *pre-group, Z* =2.023, *p* = 0.0431. Thus, hypothesis H_0^4 is rejected and the alternative hypothesis H_a^4 is accepted. Examining data for effect size demonstrated large effect, *d*= 0.88.

Table 5.5 presents means and standard deviation for *Simulation* by group.

Table 5.5

Simulation Notion for Cooling Water System

	n	M	SD
Pre	16	2.172	0.478
Post	17	2.341	0.682

Dataset met outliers, normal distribution, and homogeneity of variance requirements. Pooled t-test was applied to the dataset; results indicated no significant main effect, t=0.8207, *p*=.2090. Thus, hypothesis H_0^5 is accepted.

5.4 Discussion

Engineering and technology students may work in facilities with industrial systems. Industrial systems are complex to understand, and students may not be well prepared to perform their duties due to their complexity. When students are introduced to systems during their educational endeavors, they are in a form of a 2D schematic diagram or a 2D image of a 3D model. Developing proper foundation for understanding systems may take months. Graduates that have been properly prepared for working with industrial system may have advantage as they may be able to get engage with industrial systems soon after being hired.

The term mental model for industrial system and the level of mental model can represent individual level of understanding and consequently, the individual level of preparedness to work with industrial system. De Kleer and Brown (1983) distinguished four notions of mechanistic mental models: *Device Topology*, *Envisioning*, *Causal Model*, and *Simulation*. This work adopts the framework proposed by De Kleer and Brown, with certain adjustment to reflect on the added complexity of industrial system in comparison to devices. Specifically, the notions listed in this study are *Process System Topology*, *Service System Topology*, *Envisioning*, *Causal Model, and Simulation*. To examine the potential of a VR framework that properly prepares students to work with industrial systems, two applications were developed: CWVR, a prefabricated cooling water VR application where students can review the cooling system and interact with its components, and SDVR, an application where students can design and build industrial systems. A pool of students has been divided into two groups. The first group started their journey in this study by reviewing and interacting with the prefabricated cooling system in CWVR known as the *Pre-*

group. The *Pre-group* moved to the SDVR application after completing their experience with CWVR. The second half started their journey designing and building a system with SDVR, and then moved to reviewing and interacting with the prefabricated cooling system in CWVR known as the *Post-group*. Level of notions of mental models for both groups were assessed following their experiences with CWVR and SDVR. For the purpose of this study, the notions of the mental models of the CWVR were compared with the hope to find that experiencing SDVR before CWVR led to higher level mental model notions of the cooling water system in comparison to students that delved directly into CWVR. Should this occur, then there are evidences that interacting with VR application for industrial system has a potential to enhance student mental models of industrial systems, and consequently, properly prepare engineering and technology students to work with industrial systems in their professional lives.

Tables 5.1 and 5.2 demonstrated level of mental model notions of *Process* and *Service System Topologies*, marginally higher in the post-group. Table 5.4 shows that notion of *Causal Model* was significantly higher in the post-group. No differences in levels of notions of *Envisioning* and *Simulation* were detected between the post and pre-groups. Effect sizes for the notions of *Process* and *Services System Topology* was moderate; effect size for the notion of *Causal Model* was large. The following discusses the results above.

De Kleer and Brown (1983) distinguished among three types of learning with respect to attempting to understand mechanical systems:

1. In the first type learning, learners are developing understanding of the relationship between the components of the system and their functions (e.g., *System Topology* and *Envisioning* notions) to form fundamental knowledge of industrial system. Our results indicate that students with prior experience with systems (e.g., experience with SDVR)

were indeed able to better recognize and recall system components but their ability to understand the functions of the components did not improved significantly. That is, prior engagement with system as provided in this experiment led to partial learning described as first type by De Kleer and Brown.

2. In the second type of learning, learners convert implicit assumptions about the system to explicit, to form a stronger relationship between the coupled *system topology-envisioning* and the system function, thus forming a *Causal Model* for the system. It is possible that multiple Causal Models will be formed as some assumptions may be ambiguous or false. The results herein demonstrated that prior experience with SDVR yielded significant elevated level of *Causal Model notion*; again, the *notion* of *Causal Model* was enhanced although there was no difference on *Envisioning*.
3. In the third type of learning, the learner is utilizing a causal model to project the system's behavior, through simulating and problem solving…

   '*The third form of learning concerns a technique for preserving this "work" so that it can be called upon only when needed and otherwise remains transparent. In essence, one can cache the results of projection (namely problem-solving) by recording what aspects of the component models were actually used in the device's correct causal model'* (De Kleer & Brown, 1983, p. 183).

Learning to prepare students to meet the challenges associated with their jobs can be perceived as a dual phase process. In the first phase, the learner state of knowledge changes. In the second phase, the learner works toward mastering the knowledge gained in the first phase. This transition is referred to in the literature as 'naïve-expert shift' (Wiser & Carey, 1983). The naïve-expert shift involves conceptual change where successive concepts are different in (1) the phenomena at hand and how the phenomena deliver its purpose; (2) system model and the ability to explain to the mechanics of the system; and (3) the concept (Wiser & Carey, 1983). It is important to note that as mastery evolves, the collection of concepts may include concepts that did not exist earlier in the shift.

Returning to the results, engineering and technology students are exposed to systems in their curriculum. They were exposed to the concept of systems and it is safe to assume their experience can marked naïve. Students in the *Post-group* were asked to review a given problem statement (e.g., build a system for transporting hot water and collecting cold water, purity water) and solve the problem by designing and building a system. The engagement with SDVR forces students to generate a structure (*System Topology*), then entertain component-function relationships (*Envisioning*) to form a causal model, and then mentally simulate the system to verify it solves the problem at hand; thus engaging the students in all mental model notions. Projecting on conceptual changes through the 'naïve-expert shift', the elevated notion of *Causal Model* level in the *Post-group* can potentially be attributed to a conceptual change associated with an enhanced ability to explain mechanics of the CWVR system due to their brief exposure with SDVR. Furthermore, since both VR applications (CWVR and SDVR) involve industrial systems that change fluid temperatures, the *Post-group* had a slight advantage in developing *Process and Service System Topology* notions due to their experience in SDVR. Yet, their very limited exposure did not deliver skill enhancements that elevated levels of *Simulation* notion. It is important to note though that while *Post-group*'s *Causal Model* notion was significantly higher than *Pre-group*, the mean for the *Post-group* was only M=2.515, which is the midrange between the ratings of 'Fair' and 'Good'. On the other hand, *Pre-group's Causal Model* mean was M=1.895, which is the top range between the ratings of 'Poor' and 'Fair'. The results herein hint on the potential of VR-applications such SDVR to develop mental models of systems. It is important to note that the modest-high level of notions of *Process and Service Topology* and this of *Causal* Model occurred after a single exposure to SDVR. Examining the level of difference following multiple exposure will paint a clearer picture of the merit of an application such as SDVR.

CHAPTER 6. CAN LEVEL OF PRESENCE PREDICT LEVEL OF NOTIONS OF MENTAL MODELS IN VIRTUAL REALITY APPLICATIONS FOR DESIGNING AND BUILDING INDUSTRIAL SYSTEMS

6.1 Research Question and Hypotheses

The overarching research question is whether students' sense of *presence* in the SDVR application is associated with the students' level of notions mental models of industrial system. The overarching research question will be explored with the following five hypotheses:

Presence association with the notion of *Process System Topology*:

H_0^1: Presence does not significantly predict the outcome of *Process System Topology*.

H_a^1: Presence significantly predicts level of notion of *Process System Topology*.

Presence association with the notion of *Service System Topology*:

H_0^2: Presence does not significantly predict the outcome of *Service System Topology*.

H_a^2: Presence significantly predicts level of notion of Service *System Topology*.

Presence association with the notion of *Envisioning*:

H_0^3: Presence does not significantly predict the outcome of *Envisioning*.

H_a^3: Presence significantly predicts level of notion of *Envisioning*.

Presence association with the notion of *Causal Model*:

H_0^4: Presence does not significantly predict the outcome of *Causal Model*.

H_a^4: Presence significantly predicts level of notion of *Causal Model*.

Presence association with the notion of *Simulation*:

H_0^5: Presence does not significantly predict the outcome of *Simulation*.

H_a^5: Presence significantly predicts level of notion of *Simulation*.

Presence assessment followed protocol by Osoh et al (2000). Presence assessment instrument is referred to as SUS.

6.2 Evaluation Results

Multiple logistic regressions have been modeled for each one of the notions of mental model. While the hypotheses above address presence as a predictor, in order to shed further light on the observations from the logistics model, other data collected in the interaction log were used as predictors as well. These predictors are listed below (type of variable is provided in parenthesis):

- **SUS Means *Presence*** (nominal): Depicted as either *high* or *low* based on the average response from all six questions in SUS; students with a value of ‘6’ or ‘7’ were considered experiencing high level of *presence*; other were considered experiencing low presence category
- **SUS Count** (ordinal): Depicted as the number of student responses out of six questions that corresponded with a rating of ‘6’ or ‘7’.
- **Group** (nominal): Pre-group or Post-group, indicting application order as described in Chapter 5.
- **Interaction time** (continuous): Time students spent in SDVR since they started interacting with the component dispenser menu.
- **Overall time** (continuous): Overall time students spent in SDVR.
- **Grabbing**: Number of times students grabbed components.

- **Snapping**: Number of times students snapped components to each other, to the steam boiler, or to the water ports.
- **Process component count** (continuous): Number of process components in the student's final system.
- **Service component count** (continuous): Number of service components in the student's final system.

For 'SUS Count', 27 students ranked *low* on *presence* and six ranked *high*. This imbalanced prevented developing stable regression models, and thus SUS Count was nor included further. For Group, 16 students were in the Pre-group and 17 students in the Post-group.

All mental model notions in the multiple logistic regression analyses are nominal. Mental model notions received a binary outcome of *high level* or *not high level* as follows:

- Students that scored 80 or higher in Process System Topology, Process Service Topology, or in Envisioning are consider having a *high-level* notion; otherwise, they are not considered *high level* notion.
- Students that scored 3 are consider having a *high-level* notion or higher in the Causal Model and Simulation notions; otherwise, they are not considered having a *high-level* notion.

For the alternative hypotheses to be accepted, significant model fit and significant association with SUS Means Presence must be established.

Table 6.1 shows multiple logistic regression data for Process System Topology.

Table 6.1
Multiple Logistic Regression for Process System Topology

Predictor	β	*SE* β	χ^2	*p*
SUS Means *Presence* (0 = Low, 1 = High)	-0.3956	0.5338	0.5655	.4520
Group (0 = Post-group, 1 = Pre-group)	0.0580	0.4292	0.0183	.8923
Interaction Time	0.0010	0.0049	0.0384	.8447
Overall Time	-0.0010	0.0046	0.0437	.8344
Grabbing	-0.0062	0.0067	0.938	.3329
Snapping	0.0123	0.0124	1.0543	.3045
Process Component Count	-0.0835	0.0949	0.8356	.3607
Service Component Count	0.1509	0.1518	1.0097	.3150
Likelihood Ratio Chi-square Test (Whole Model Test)	-	-	2.4601	.9636

Note. R^2 = 0.0547

Results from the likelihood ratio chi-square test indicates the logistic model does not significantly fit better than the null model only, $\chi^2(8, N = 33) = 2.4601$, $p = .9636$; no significant association between any of the predictor and Process System Topology were detected, thus hypothesis H_0^1 is accepted.

Table 6.2 shows the multiple logistic regression data for Service System Topology.

Table 6.2
Multiple Logistic Regression for Service System Topology

Predictor	β	*SE* β	χ^2	*p*
SUS Means *Presence* (0 = Low, 1 = High)	-3.3928	1.4667	11.1653	0.0008^{**}
Group (0 = Post-group, 1 = Pre-group)	-0.9025	0.7125	1.9062	0.1674
Interaction Time	-0.0365	0.0147	13.6235	0.0002^{**}
Overall Time	0.0341	0.0138	13.2991	0.0003^{**}
Grabbing	0.0050	0.0139	0.1394	0.7089
Snapping	0.0147	0.0244	0.3834	0.5358
Process Component Count	-0.2729	0.2015	1.9089	0.1671
Service Component Count	-0.4829	0.3047	3.6322	0.0567
Likelihood Ratio Chi-square Test (Whole Model Test)	-	-	22.6872	0.0038^{*}

Note. R^2 = 0.4962, *Significant at p < .01, **Significant at p < .001

Table 6.2 shows that the logistic model fits significantly better than the null model from the likelihood ratio chi-square test, $\chi^2(8, N = 33) = 22.69$, $p = .0038$. Multiple logistic regression also shows significant association between *SUS Means Presence*, *interaction*, and *overall time*

, and Service System Topology, all with $p < .001$. Due to significant model fit and significant association with SUS Means *Presence*, the hypothesis H_0^2 is rejected and the alternative hypothesis H_a^2 is accepted.

Table 6.3 shows the multiple logistic regression data for Envisioning.

Table 6.3
Multiple Logistic Regression for Envisioning

Predictor	β	*SE* β	χ^2	*p*
SUS Means *Presence* (0 = Low, 1 = High)	0.1769	0.6633	0.0731	.7869
Group (0 = Post-group, 1 = Pre-group)	0.0306	0.5139	0.0035	.9525
Interaction Time	0.0037	0.0066	0.3483	.5551
Overall Time	-0.0038	0.0061	0.4066	.5237
Grabbing	-0.0059	0.0089	0.5186	.4714
Snapping	0.0007	0.0153	0.0024	.9612
Process Component Count	-0.0283	0.1106	0.0675	.7950
Service Component Count	-0.0021	0.1934	0.0001	.9913
Likelihood Ratio Chi-square Test (Whole Model Test)	-	-	4.1887	.8397

Note. R^2 = 0.1083

Based on the likelihood ratio chi-square test, the logistic model for Envisioning does not significantly fit better than the null model, $\chi^2(8, N = 33) = 4.1887$, $p = .8397$. The logistic model for Envisioning shows no significant association between predictors and Envisioning. Thus, hypothesis H_0^3 is accepted.

Table 6.4 shows the multiple logistic regression data for Causal Model.

Table 6.4
Multiple Logistic Regression for Causal Model

Predictor	β	*SE* β	χ^2	*p*
SUS Means *Presence* (0 = Low, 1 = High)	-0.1653	0.5978	0.0762	.7826
Group (0 = Post-group, 1 = Pre-group)	-0.5374	0.4944	1.2409	.2653
Interaction Time	-0.0061	0.0057	1.2304	.2673
Overall Time	0.0052	0.0054	1.0261	.3111
Grabbing	0.0012	0.0072	0.0275	.8684
Snapping	-0.0100	0.0146	0.4802	.4883
Process Component Count	0.2012	0.1356	3.1514	.0759
Service Component Count	-0.1725	0.1735	1.0232	.3118
Likelihood Ratio Chi-square Test (Whole Model Test)	-	-	9.3330	.3150

Note. R^2 = 0.2052

Likelihood ratio chi-square test indicates that the logistic model for Causal Model does significantly fit better than the null model only, $\chi^2(8, N = 33) = 9.333$, $p = .3150$. The logistic model for Causal Model revealed no significant association with the predictors, and thus, hypothesis H_0^4 is accepted.

Table 6.5 shows the multiple logistic regression data for Simulation.

Table 6.5
Multiple Logistic Regression for Simulation

Predictor	β	*SE* β	χ^2	*p*
SUS Means *Presence* (0 = Low, 1 = High)	-1.2445	1.0049	1.9044	0.1676
Group (0 = Post-group, 1 = Pre-group)	1.6898	0.9559	4.4613	0.0347*
Interaction Time	0.0059	0.0166	0.1264	0.7221
Overall Time	-0.0090	0.0158	0.3443	0.5573
Grabbing	0.0147	0.0137	1.7730	0.1830
Snapping	-0.0335	0.0245	2.7099	0.0997
Process Component Count	0.4791	0.2535	8.6495	0.0033**
Service Component Count	0.7398	0.4244	5.3837	0.0203*
Likelihood Ratio Chi-square Test (Whole Model Test)	-	-	21.5080	.0059

Note. $R^2 = 0.5313$, *Significant at p < .05, **Significant at p < .01

Based on the likelihood ratio chi-square test, $\chi^2(8, N = 33) = 21.51$, $p = .0059$ in table 6.5, the logistic model fits significantly better than the null model. Simulation is significantly associated with Group, System Process Count, and Service Component counts. The Post-group has higher log odds for Strong Simulation notion than the Pre-group. Despite the significant model fit, there was no significant association with *SUS Means Presence*, thus hypothesis H_0^5 is accepted.

6.3 Discussion

Analysis of data presented in Tables 6.1-6.5 revealed that *presence* is a significant predictor of high-level of notion only for Service System Topology. The following section will explore the finding with *presence*. Further, the following discussion attempt to establish the role of time factors, interactions, and component count with *presence*.

Tables 6.1-6.5 reveal significant associations between *presence* level and other factors and their respective mental model notions. Service System Topology was the only mental model notion that had a significant model fit and was significantly associated with level of *presence*; that is, students marked *high-level* of *presence* had a *high-level* notion of Service System Topology. Literature review failed short in identifying significant research on the relationships between *Presence* and mental representations. In one study, (Mikropoulos, 2006) *presence* was documented to helped users perform their learning tasks successfully and to positively contribute to learning outcomes. Further, the study found that users representation model played important role in their sense of self-presence. However, the study herein is investigating the opposite relationship, e.g., the extent to which *presence* can predict mental representation. Bogicevic and colleagues (2019) examined the role of mental imagery on inducing presence in VR in tourism. Their results indicate that interactive VR platform induced high elaboration of mental imagery and higher sense of presence. Additional study presents that presence in VR lead to positive change in attitude toward destination in tourism and that VR is more persuasive when the VE conveys situated affordances (Tussyadiah, 2018). As mentioned earlier, published research on the relationships between *Presence* and mental models revealed little significant information.

Since Service System Topology represents the understanding of the system's service components, students become preoccupied when working with service components, potentially elevating the students' mental workload to establish the notion of Service System Topology. Ma and Kaber (2006) provided evidences for significant relationship between presence and workload but could not support a correlation between presence and performance. They also state that their results have general applicability for the design of multimodal interfaces for real-world task category, which the study herein falls under. The component menu and component handling in this study represents high level interactivity. Interactivity related factors herein include interaction time, number of grabs and snaps, and number of process and service components students implemented during designing and building their system. Appendix B consists of detailed results of multiple logistic regression of each one of the questions in SUS with each of the factors discussed earlier. Table 6.6 presents selected data from these models. The data includes significance level from effect likelihood ratio tests, R^2, and significance level from likelihood ratio chi-square test. Individual SUS Questions are treated as nominal data where ratings of '6' or '7' are considered *high-level* presence.

Table 6.6. Multiple Logistic Regression on SUS: Data from Effect Likelihood Ratio Tests

Predictor	Q1	Q2	Q3	Q4	Q5	Q6	SUS Means *Presence*
Group	.0286*	.6238	.2573	.7585	.4289	.4000	.8068
Overall Time (sec)	.8765	.0717**	.4800	.9150	.4506	.0903**	.6227
Interaction Time (sec)	.9600	.0949**	.3606	.9098	.5736	.1275	.7504
Number of Snaps	.6314	.0481*	.8008	.2536	.0156*	.0241*	.0508**
Number of Grabs	.8940	.0636**	.1346	.0753**	.0662**	.1139	.2043
Process Component Count	.7567	.0254*	.5705	.0442*	.5226	.3891	.3520
Service Component Count	.0475*	.8890	.0165*	.4944	.2492	.6620	.6379
R-square (R^2)	.2709	.4156	.3256	.2747	.1919	.3116	.2027
Likelihood Ratio Chi-square Test (Whole Model Test)	.0905**	.0147*	.0678**	.0894**	.3864	.0905**	.5004

*Significance at $p < 0.05$, **Marginal significance at $p < 0.1$

The *number of snaps* factor is significant associated with presence questions 2, 5 and 6, and has marginal significant associations with overall *presence*. *Number of Grabs* has marginal significant associations with presence questions 2, 4 and 5. Since *Number of Grabs* and *number of snaps*, the interactive elements in SDVR, are likely to induce enhanced workload, and workload has significant relationship with presence (Ma & Kaber, 2006), one may conclude that there is a causal chain of association between these predictors and the overall sense of *presence* (SUS Means). When observing Table 6.6, only *number of snaps* has significant association with *presence* and this significance is marginal. Returning to the relationship between notion of Service System Topology, and *presence* and the other factors, *presence* was a significant predictor of Service System Topology, but neither *number of grabs* nor *number of snaps* were. Therefore, based on the indirect relationships observed through Tables 6.1 and 6.6, there is a potential for case for a causal relationship between interactivity level and notion of System Service Topology. One important fact that may have a significant impact on the results is that many students rated *not high-level* on presence. Potentially, low level presence may prevent establishing relationships with the notions of mental models. To get further perspective of this claim, the number of students that rated *high-level on* each of SUS presence questions are presented in Figure 1. As can be observed from the figure, numbers of *high-level* ratings are disappointing.

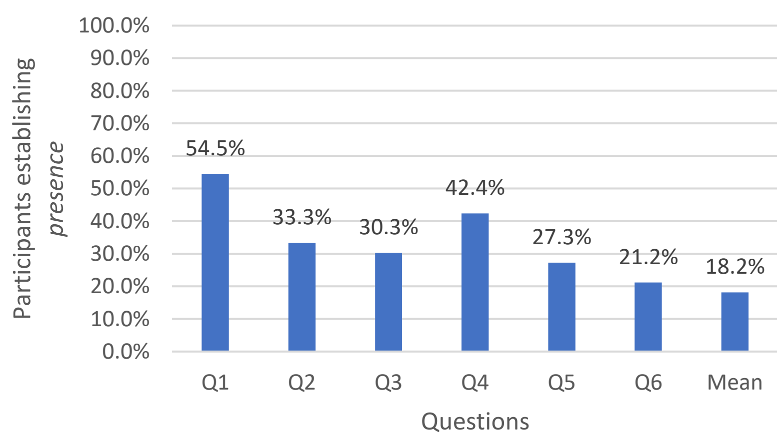

*Figure 6.1***.** High-level presence count by presence questions.

The low numbers on questions 1, 2, 4, and 6 and the overall mean for presence may be attributed to break-in-*presence*. Members of the research team sometimes needed to alert students to prevent them from walking into walls, colliding with objects, or with each other. Further, Copper et al (2018) presents that high-level *presence* is a function of channels of feedback in the environment (e.g., audio, visual, and vibrotactile feedback). In SDVR, while designed for establishing a strong sense of agency, the only other feedback was visual. Adding credible ambient sound of industrial environment may help. Furthermore, SVDR did not include self-representation. Self-representation is reported as a significant contributor of personal *presence* (e.g., Mikroupolus, 2006; Pritchard et al, 2016). Even further, implementing some haptic tactile though the controllers when users are interacting with components could assist with enhancing *presence*

The numbers of *high-level* in questions 3 and 5 may be attributed to the fact that most students did not visit a machine room physically, nevertheless on a daily basis. Since these questions examine presence with respect to experience in reality, students did not have a point of reference, thus their rating maybe irrelevant.

In summary, presence was a significant predictor for notion Service System Topology only. Low levels of presence may be a concern as they might prevented positive learning outcomes with other notions of mental models. Potentially, **break in presence** described above was a significant hindering factor. Similarly, lack of self- representation and limited channels of feedback reduced the likelihood of establishing strong presence. Moreover, aspects of frustration, workload, temporal constraints, that have not been measured could be inhibiting factors.

CHAPTER 7. SUMMARY & CONCLUSION

To properly work with industrial systems, engineering and technology students must have the fundamental understanding of industrial systems. Due to their complexity, industrial systems can be a challenging subject and students may not be well prepared to work with industrial systems. This study explored the following two questions:

1. *Can VR enhance students' mental models of industrial systems?*
2. *Does presence contribute to the enhancement of students' mental models of industrial systems?*

Two VR applications, CWVR and SDVR, were developed for assessing level of mental model notions in students. Student were divided two groups, pre and post: the Pre-group started with a task with CWVR and then moved to work on a task with SDVR, and vice versa for the Post-group. Student mental models were assessed following each task. Additionally, students' presence was also assessed with the SUS instrument after their engagement with SDVR.

The results point to a potential merit with using VR application such as SDVR to develop mental model in students. Further, *presence* was found as significant predictor of one of the notions of mental models, *Service System Topology*.

The conclusions herein are that the fact that a single exposure with the SDVR application yielded positive enhancement in notions of mental models, then utilizing an application such as SDVR for multiple sessions in college curricula may have merit in the journey to create robust systems mental models in technology and engineering students. Additionally, the results on the relationships between *presence* and notion of mental models, while significant with some factors, can be determined as causal at best. Further research will require using a different presence instrument and examining other factors such as cognitive indicators in order to explain the results.

7.1 Limitations

Conducting the VR experiments herein is demanding. The development of the VR applications is an effortful and tedious process. In the study, a complete experiment session with a single user took approximately 90 minutes. Student average time in SDVR only was approximately 44 minutes long. Thus, conducting the study with a large number of participants is almost not feasible. Also, several students commented that the problem task in SDVR was complex.

The small sample size also prevented incorporating SUS Presence Count in the analysis since the low count did not allow fitting a model with a logistic regression. with SUS Count, the presence metric of SUS Questionnaires, as the predictor and response due to small observations reported in the higher levels of presence. Finally, the use of SUS for measuring *presence* hindered performing a more thorough investigation into the relationship between presence in SDVR and notions of mental models of systems.

REFERENCES

Alhalabi, W. S. (2016). Virtual reality systems enhance students' achievements in engineering education. *Behaviour & Information Technology*, DOI:10.1080/0144929X.2016.1212931

Austin, L. C., Kovacs, D., Throne, S., Moody, J. R.K. (2020, *accepted*). Using grounded theory and mental modeling to understand influences on electricians' safety decisions: Toward an integrated theory of why electricians work energized. *Safety Science*, 130.

Bogicevic, V., Seo, S., Kandampully, J. A., Liu, S. Q., & Rudd, N. A. (2019). Virtual reality presence as a preamble of tourism experience: The role of mental imagery. *Tourism Management, 74,* 55-64. doi:10.1016/j.tourman.2019.02.009

Brelsford, J. W. (1993). Physics Education in a Virtual Environment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 37(18), 1286–1290. https://doi.org/10.1177/154193129303701818

Brooks, Jr., F. P. (1999). What's Real About Virtual Reality? In *IEEE Computer Graphics and Applications,* 19(6), 16-27. https://doi.org/10.1109/38.799723

Chen. S, She, H., & Hsiao, M. (2015). Using Eye-Tracking to Investigate the Different 3D Representation on Students' Mental Model Construction. *2015 IEEE 15th International Conference on Advanced Learning Technologies*. 388-390.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences,* Lawrence Erlbaums Associates, New York: NY.

Cox, P., Niewohner, J., Pidgeon, N., Gerrard, S., Fischhoff, B., & Riley, D. (2003). The Use of Mental Models in Chemical Risk Protection: Developing a Generic Workplace Methodology. *Risk Analysis*, 23(2), 311-324.

Cruz-Neira, C. Sandin, D. J., DeFanti, T. A. (1993). Surround-screen projection-based virtual reality: the design and implementation of the CAVE, in SINGRAPH '93 *proceeding of the 20th annual conference on Computer graphics and interactive techniques,* pp. 135-142.

De Kleer, J. & Brown, J. S. (1983). Assumptions and ambiguities in mechanistic models. In Gentner, D & Stevens, A. L. (Eds.), *Mental Models*, 155-190.

Dinis, F. M., Guimaraes, A. S., Carvalho, B. R., & Martins, J. P. P. (2017). Virtual and augmented reality game-based applications to civil engineering education. In *2017 IEEE Global Engineering Education Conference (EDUCON)*, April 25-28, Athens, Greece. https://doi.org/10.1109/EDUCON.2017.7943075

Furlough, C. S. & Gillan, D. J. (2018). Mental Models: Structural Differences and the Role of Experience. *Journal of Cognitive Engineering and Decision Making*, 12(4), 269-287.

Harackiewicz, J. M., & Knogler, M. (2017). Interest: Theory and application. In A. J. Elliot, C. S. Dweck, & D. S. Yeager (Eds.), *Handbook of competence and motivation: Theory and application* (p. 334–352). The Guilford Press.

Havig, P., McIntire, J., & Geiselman, E. (2011). Virtual reality in a cave: limitations and the need for HMDs? In *Proc. SPIE 8041*, *Head- and Helmet Mounted Displays XVI: Design and Applications*, 804107, June 1, Orlando, Florida, USA. https://doi.org/10.1117/12.883855

Hegarty, M., Stieff, M., & Dixon, B. L. (2013). Cognitive change in mental models with experience in the domain of organic chemistry. *Journal of Cognitive Psychology*, 25(2), 220-228.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, 30, 141-158.

Holman, L. (2011). Millennial Students' Mental Models of Search: Implications for Academic Librarians and Database Developers. *The Journal of Academic Librarianship*, 31(1), 19-27.

Horan, B., Sevedmahmoudian, M., Mortimer, M., Thirunavukkarasu, G. S., Smilevski, S., & Stojcevski, A. (2018). Feeling Your Way Around a CAVE-Like Reconfigurable VR System. In *2018 11th International Conference on Human system Interaction (HIS)*, July 4-6, Gdansk, Poland.

HTC Corporation. (n. d.). VIVE | Discover Virtual Reality Beyond Imagination. Retrieved from http://www.vive.com.

Jalmo, T. & Suwandi, T. (2018). Biology education students' mental models on genetic concepts. *Journal of Baltic Science Education*, 17(3), 474-485.

Kang, R., Dabbish, L., Fruchter, N., & Sara Kiesler. (2015). My Data Just Goes Everywhere: User Mental Models of the Internet and Implications for Privacy and Security. In *Symposium on Usable Privacy and Security*. Ottawa, CA.

Lackey, S. J., Salcedo, J. N., Szalma, J. L., & Hancock, P. A. (2016). The stress and workload of virtual reality training: the effects of *presence*, immersion and flow. *Ergonomics*, *59*(8), 1060-1072, DOI: 10.1080/00140139.2015.1122234

Lau, W. W.F. & Yuen, A. H.K. (2010). Promoting conceptual change of learning sorting algorithm through the diagnosis of mental models: The effects of gender and learning styles. *Computer & Education*, 54(1), 275-288.

Ma, R., & Kaber, D. B. (2006). Presence, workload and performance effects of synthetic environment design factors. *International Journal of Human-Computer Studies, 64*(6), 541-552. doi:10.1016/j.ijhcs.2005.12.003

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, 60, 225-236.

Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning*, 33, 403-423.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive Constraints on Multimedia Learning: When Presenting More Material Results in Less Understanding. *Journal of Educational Psychology*, 93(1), 187-198. http://dx.doi.org/10.103/0022-0663.93.1.187

Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology*, *100*(2), 380–386. https://doi.org/10.1037/0022-0663.100.2.380

Mayer, R. E. & Moreno, R. (1998). A Split-Attention Effect in Multimedia Learning: Evidence for Dual Processing Systems in Working Memory. *Journal of Educational Psychology*, 90(2), 312-320.

Mikropoulos, T. A. (2006). Presence: a unique characteristic in educational virtual environments. *Virtual Reality, 10*(3-4), 197-206.

Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education, 56*(3), 769-780. doi:10.1016/j.compedu.2010.10.020

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1994). Augmented reality: a class of displays on the reality-virtuality continuum. In *SPIE 2351 Telemanipulator and Telepresence Technologies*, December 21, Boston, MA.

Moreno, R. & Mayer, R. E. (1999). Cognitive Principles of Multimedia Learning: The Role of Modality and Contiguity. *Journal of Educational Psychology*, 91(2), 358-368.

Noordzij, M. L, Zuidhoek, S., & Postma, A. (2006). The influence of visual experience on the ability to form spatial mental models based on route and survey descriptions. *Cognition*, 100(2), 321-342.

Norman, D. (1983). Some Observations on Mental Models. *Mental Models*, 8-14.

Norman, D. (2013). *The Design of Everyday Things*, Basic Books, New York: NY.

Oculus, Inc. (n. d.). Oculus | VR Headset & Equipment. Retrieved from http://www.oculus.com.

Pantförder, D., Schaupp, J., & Vogel-Heuser, B. (2017). Making Implicit Knowledge Explicit – Acquisition of Plant Staff's Mental Models as a Basis for Developing a Decision Support System. In Stephanidis C. (Eds.), *HCI International 2017 – Posters' Extended Abstracts. HCI 2017. Communications in Computer and Information Science*, 713, Springer, Cham.

Park, B., Flowerday, T., & Brünken, R. (2015). Cognitive and affective effects of seductive details in multimedia learning. *Computers in Human Behavior*, 44, 267-278.

Parong, J. & Mayer, R. E. (2018). Learning Science in Immersive Virtual Reality. *Journal of Educational Psychology*, 110(6), 785-797. http://dx.doi.org/10.1037/edu00024

Patwardhan, M. & Murthy, S. (2015). When does higher degree of interaction lead to higher learning in visualizations? Exploring the role of 'Interactivity Enriching Features'. *Computers & Education*, 82, 292-305.

Puschmann, P., Horlitz, T., Wittstock, V., & Schutz, A. (2016). Risk Analysis (Assessment) Using Virtual Reality Technology – Effects of Subjective Experience: An Experimental Study. In *Procedia CIRP*, 50, 490-495. https://doi.org/10.1016/j.procir.2016.04.115

Pritchard, S. C., Zopf, R., Polito, V., Kaplan, D. M., & Williams, M. A. (2016). Non-hierarchical Influence of Visual Form, Touch, and Position Cues on Embodiment, Agency, and Presence in Virtual Reality. *Frontiers in Psychology, 7*. doi:10.3389/fpsyg.2016.01649

Revell, K. M. A. & Stanton, N. A. (2014). Case studies of mental models in home heat control: Searching for feedback valve, timer, and switch theories. *Applied Ergonomics*, 45(3), 363-378.

Rowe, A. L. & Cooke, N. J. (1995). Measuring Mental Models: Choosing the Right Tools for the Job. *Human Resource Development Quarterly*, 6(3), 243-255.

Rupp, M. A., Odette, K. L., Kozachuk, J., Michaelis, J. R., Smither, J. A., & McConnell, D. S. (2019). Investigating learning outcomes and subjective experiences in 360-degree videos. *Computers & Education*, 128, 256-268. https://doi.org/10.1016/j.compedu.2018.09.015

SAS Institute, Inc. (n. d.). Statistical Software | JMP Software from SAS. Retrieved July 26, 2020, from https://www.jmp.com/en_us/home.html

Schunk, D. H. & Pajares F. (2009). Self-Efficacy Theory. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of Motivation at School* (pp. 36-53), New York, NY: Taylor & Francis

Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (p. 183–212). Lawrence Erlbaum Associates, Inc.

Tussyadiah, I. P., Wang, D., Jung, T. H., & Dieck, M. T. (2018). Virtual reality, presence, and attitude change: Empirical evidence from tourism. *Tourism Management, 66*, 140-154. doi:10.1016/j.tourman.2017.12.003

Valve Software. (n. d). Valve Index – Upgrade Your Experience. Retrieved from https://www.valvesoftware.com/en/index.

Wang, F., Li, W., Mayer, R. E., & Liu, H. (2018). Animated pedagogical agents as aids in multimedia learning: Effects on eye-fixations during learning and learning outcomes. *Journal of Educational Psychology*, 110(2), 250-268. http://dx.doi.org/10.1037/edu0000221

Webster, R. (2015). Declarative knowledge acquisition in immersive virtual learning environments. *Interactive Learning Environments*, 24(6), 1319-1333.

Wiser, M. & Carey, S. (1983). When heat and temperature were one. In Gentner, D & Stevens, A. L. (Eds), *Mental Models*, 267-297.

Wyk, E. &Villiers, R. (2009). Virtual Reality Training Applications for the Mining Industry. In *AFRIGRAPH '09 Proceedings of the 6th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa*, 53-63. https://doi.org/10.1145/1503454.1503465

APPENDIX A. IRB 19-607

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Institutional Review Board
Office for Responsible Research
Vice President for Research
2420 Lincoln Way, Suite 202
Ames, Iowa 50014
515 294-4566

Date: 10/30/2019

To: Robert Slezak Nir Keren

From: Office for Responsible Research

Title: **Research Study about Enhancing User Mental Models of Industrial Systems in Virtual Reality**

IRB ID: **19-507**

Submission Type: Initial Submission **Exemption Date:** 10/30/2019

The project referenced above has been declared exempt from most requirements of the human subject protections regulations as described in 45 CFR 46.104 or 21 CFR 56.104 because it meets the following federal requirements for exemption:

2018 - 1: Research, conducted in established or commonly accepted educational settings, that specifically involves normal educational practices that are not likely to adversely impact students' opportunity to learn required educational content or the assessment of educators who provide instruction. This includes most research on regular and special education instructional strategies, and research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.

The determination of exemption means that:

- **You do not need to submit an application for continuing review. Instead, you will receive a request for a brief status update every three years. The status update is intended to verify that the study is still ongoing.**
- **You must carry out the research as described in the IRB application.** Review by IRB staff is required prior to implementing modifications that may change the exempt status of the research. In general, review is required for any *modifications to the research procedures* (e.g., method of data collection, nature or scope of information to be collected, nature or duration of behavioral interventions, use of deception, etc.), any change in *privacy or confidentiality protections*, modifications that result in the *inclusion of participants from vulnerable populations*, removing plans for informing participants about the study, any *change that may increase the risk or discomfort to participants, and/or* any change such that the revised procedures do not fall into one or more of the regulatory exemption categories. The purpose of review is to determine if the project still meets the federal criteria for exemption.
- All ***changes to key personnel*** must receive prior approval.
- **Promptly inform the IRB of any addition of or change in federal funding for this study.** Approval of the protocol referenced above applies only to funding sources that are specifically identified in the corresponding IRB application.

APPENDIX B. COOLING WATER ASSESSMENT RUBRIC FOR CAUSAL MODEL AND SIMULATION NOTIONS

	Excellent	Good	Fair	Poor
Causal Model Causal model describes the overall function of the system in terms of how the components interact.	The user's causal model of the cooling water system is excellent. The user's response shows general understanding of the system and mentions all important components that contribute to the function.	The user's causal model of the cooling water system is good. The user's response shows understanding of the system's function of the system and mentions most of the important components that contribute to the function.	The user's causal model of the cooling water system is fair. The user's response shows partial understanding of the system's function. Some important components that contribute to the system function are mentioned.	The user's causal model of the cooling water system is poor. The user's response shows lack of understanding the cooling water system's function or gave an irrelevant response. None to few important components are mentioned.
Simulation Simulation refers to running the system to produce specific behavior.	The simulation of the cooling water system is excellent. The user's response shows a detailed understanding of how the system behaves.	The simulation of the cooling water system is good. The user's response shows general understanding of how the system behaves. The response is almost complete.	The simulation of the cooling water system is fair. The user's response identifies how the system will behave but does not expand onto the implications of the behavior.	The simulation of the cooling water system is poor. The user's response shows lack of understanding on how the cooling water system will behave or gave an irrelevant response.

APPENDIX C. MULTIPLE LOGISTIC REGRESSION TABLES FOR SUS QUESTIONS AND MEANS PRESENCE

Multiple Logistic Regression for SUS Question 1

Predictor	β	*SE* β	χ^2	*p*
Group (0 = Post-group, 1 = Pre-group)	1.0944	0.5484	4.7898	.0286*
Interaction Time	0.0003	0.0059	0.0025	.9600
Overall Time	0.0008	0.0055	0.0241	.8765
Grabbing	-0.0010	0.0076	0.0177	.8940
Snapping	-0.0066	0.0140	0.2301	.6314
Process Component Count	-0.0354	0.1120	0.0960	.7567
Service Component Count	0.3755	0.2166	3.9270	.0475*
Likelihood Ratio Chi-square Test (Whole Model Test)	-	-	12.3203	.0905

Note. R^2 = 0.2709, *Significant at p < .05

Multiple Logistic Regression for SUS Question 2

Predictor	β	*SE* β	χ^2	*p*
Group (0 = Post-group, 1 = Pre-group)	-0.2989	0.6135	0.2405	.6238
Interaction Time	-0.0136	0.0101	2.7899	.0949
Overall Time	0.0135	0.0095	3.2427	.0717
Grabbing	0.0156	0.0094	3.4417	.0636
Snapping	-0.0298	0.0174	3.9071	.0481*
Process Component Count	0.3810	0.2478	4.9938	.0254*
Service Component Count	0.0287	0.2059	0.0195	.8890
Likelihood Ratio Chi-square Test (Whole Model Test)	-	-	17.4613	.0147*

Note. R^2 = 0.4156, *Significant at p < .05

Multiple Logistic Regression for SUS Question 3

Predictor	β	*SE* β	χ^2	*p*
Group (0 = Post-group, 1 = Pre-group)	0.6738	0.6196	1.2834	.2573
Interaction Time	0.0065	0.0071	0.8358	.3606
Overall Time	-0.0045	0.0064	0.4988	.4800
Grabbing	-0.0159	0.0140	2.2384	.1346
Snapping	0.0049	0.0199	0.0636	.8008
Process Component Count	-0.0625	0.1109	0.3218	.5705
Service Component Count	0.4810	0.2365	5.7503	.0165*
Likelihood Ratio Chi-square Test (Whole Model Test)	-	-	13.1828	.0678

Note. R^2 = 0.3256, *Significant at p < .05

Multiple Logistic Regression for SUS Question 4

Predictor	β	*SE* β	χ^2	*p*
Group (0 = Post-group, 1 = Pre-group)	-0.1526	0.4963	0.0945	.7585
Interaction Time	0.0008	0.0067	0.0128	.9098
Overall Time	-0.0007	0.0062	0.0114	.9150

Grabbing	0.0123	0.0081	3.1633	.0753
Snapping	-0.0145	0.0135	1.3034	.2536
Process Component Count	0.2170	0.1364	4.0499	.0442*
Service Component Count	-0.1214	0.1813	0.4670	.4944
Likelihood Ratio Chi-square Test (Whole Model Test)	-	-	12.3573	.0894

Note. R^2 = 0.2742, *Significant at p < .05

Multiple Logistic Regression for SUS Question 5

Predictor	β	*SE* β	χ^2	*p*
Group (0 = Post-group, 1 = Pre-group)	-0.4156	0.5317	0.6259	.4289
Interaction Time	-0.0033	0.0058	0.3122	.5763
Overall Time	0.0041	0.0054	0.5691	.4506
Grabbing	0.0145	0.0096	3.3759	.0662
Snapping	-0.0336	0.0174	5.8520	.0156*
Process Component Count	0.0626	0.0998	0.4087	.5226
Service Component Count	-0.2042	0.1849	1.3276	.2492
Likelihood Ratio Chi-square Test (Whole Model Test)	-	-	7.4206	.3864

Note. R^2 = 0.1919, *Significant at p < .05

Multiple Logistic Regression for SUS Question 6

Predictor	β	*SE* β	χ^2	*p*
Group (0 = Post-group, 1 = Pre-group)	0.5542	0.6913	0.7082	.4000
Interaction Time	-0.0122	0.0085	2.3225	.1275
Overall Time	0.0123	0.0078	2.8688	.0903
Grabbing	0.0139	0.0104	2.4986	.1139
Snapping	-0.0380	0.0208	5.0886	.0241*
Process Component Count	0.1080	0.1398	0.7416	.3891
Service Component Count	-0.1051	0.2434	0.1911	.6620
Likelihood Ratio Chi-square Test (Whole Model Test)	-	-	10.6289	.1556

Note. R^2 = 0.3116, *Significant at p < .05

Multiple Logistic Regression for SUS Means Presence

Predictor	β	*SE* β	χ^2	*p*
Group (0 = Post-group, 1 = Pre-group)	-0.1504	0.6153	0.0598	.8068
Interaction Time	-0.0023	0.0072	0.1012	.7504
Overall Time	0.0033	0.0065	0.2421	.6227
Grabbing	0.0104	0.0085	1.6116	.2043
Snapping	-0.0291	0.0166	3.8150	.0508
Process Component Count	0.0996	0.1122	0.8663	.3520
Service Component Count	-0.0945	0.2037	0.2215	.6379
Likelihood Ratio Chi-square Test (Whole Model Test)	-	-	6.3420	.5004

Note. R^2 = 0.2027