International Journal of Engineering Education Vol. 36, No. 6, pp. 1938–1955, 2020
Printed in Great Britain
0949-149X/91 $3.00+0.00

Assessment of the Efficacy and Effectiveness of Virtual Reality Teaching Module: A Gender-Based Comparison*

VIDANELAGE L. DAYARATHNA[1], SOFIA KARAM[1], RAED JARADAT[1],**
MICHAEL A. HAMILTON[2], MORTEZA NAGAHI[1], SAYALI JOSHI[1], JUNFENG MA[1],
OMAR ASHOUR[3] and BOUTEINA DRIOUCHE[4]
[1] Department of Industrial and Systems Engineering, Mississippi State University, PO Box 9542, Mississippi State, 39762, USA.
E-mail: vld66@msstate.edu, sk1867@msstate.edu, jaradat@ise.msstate.edu, sgj74@msstate.edu, ma@ise.msstate.edu, mn852@msstate.edu
[2] Institute for Systems Engineering Research, Mississippi State University ISER, 3909 Halls Ferry Road, Vicksburg, MS, 39180, USA.
E-mail: michaelh@iser.msstate.edu
[3] Industrial Engineering Department, The Pennsylvania State University, The Behrend College, 5350 Technology Drive, Erie, PA, 16510, USA. E-mail: oma110@psu.edu
[4] Department of Aerospace Engineering, Mississippi State University, PO Box 9549, Mississippi State, 39762, USA.
E-mail: bd981@msstate.edu

The concepts and topics of manufacturing systems design and analysis are usually taught using traditional lecturing, in-class problem solving, and project-based approaches. These concepts are not easy to grasp and can be tedious when taught by traditional methods. This study presents an innovative virtual reality (VR) based approach to teach manufacturing systems concepts. To illustrate the efficacy and effectiveness of VR technology in enhancing students learning concepts, a VR queuing theory teaching module is developed. The efficacy and effectiveness of the VR module are then analyzed for male and female participants to investigate the impact of the VR environment on female engineers in science, technology, engineering, and mathematics (STEM). Simulation sickness, system usability, and user experience tools were used to assess the efficacy of the VR module, and the queuing theory quiz, NASA TLX assessment, and post-motivation measures were applied to evaluate the effectiveness of the developed VR module. Both males and females indicated higher user satisfaction in terms of system usability. Female participants perceived higher user experience than their male counterparts. Both male and female participants experienced similar simulation sickness symptoms throughout the study. The quiz score indicated that students performed well in the conceptual section for both genders. The NASA TLX results suggested that participants required low perceived work effort in regard to performing the tasks in the module. The post motivation results confirmed that the VR module created positive motivation in learning the queueing theory for both male and female students. Overall, the efficacy and effectiveness measures affirm that both male and female participants perceived a similar experience in the developed VR teaching module.

Keywords: manufacturing systems; virtual reality; engineering education; underrepresentation of women; efficacy and effectiveness

1. Introduction

In recent years, Virtual Reality (VR) has gained a massive amount of attention [1]. VR is characterized by the immersion of participation in a synthetic environment rather than the external observation of such an environment [1]. It relies on stereoscopic, three-dimensional, hand/body tracking, head-tracked displays, and binaural sound. VR has the potential to provide its users with additional features through improved perceptual fidelity. It can also boost user efficiency by reducing the cognitive load in the completion of a task. VR may enhance the quality of life in hazardous or unpleasant working conditions and will ultimately have an effect on society as a whole. Considering all these facts, it can be stated that virtual reality might be beneficial to our everyday activities. Being a universal technology, it can be extended to any domain activities. Many fields could certainly use VR more than others, and education is one such field.

Pantelidis [2] summarized several motivations for implementing VR in education, such as sophisticated visualization, collaboration, and interaction. Mikropoulos and Natsis [3] showed the ability of VR as a tool that facilitates students' understanding of the material and minimizes confusion. VR technologies are proven to be opportune and powerful in education due to their abilities to engage individuals in an immersive simulated environment. VR provides users a real-time visualization of situations and interactions with objects, which overpowers the traditional teaching tools [4–6]. With the consideration of VR's benefits in education, the authors have proposed a VR-based module (Phase One) approach to teaching manufacturing systems concepts, i.e., queueing theory concepts [7]. The results revealed that the proposed VR teaching module

** Corresponding author.

* Accepted 6 June 2020.

outshined the existing educational pedagogics regarding students' knowledge gain and level of motivation. Specifically, the purpose of this study (Phase Two) is to investigate the efficacy and effectiveness differences between male and female students when using a VR teaching approach

Numerous studies show that female engineers positively influence and benefit the U.S. economy [8]. The positive impact of workplace diversity has led to several initiatives to enhance preparation and to increase participation of individuals, particularly those who have traditionally been under-represented in science, technology, engineering, and mathematics (STEM) enterprises [9, 10]. Although women have made several contributions in almost every field in modern and historical times, the low number of women in STEM fields is a major concern to date. In a recent survey [11] conducted by the National Center for Science and Engineering Statistics (NCSES), women represented 50% of the U.S. population aged between 18 and 64. Today, more women are enrolled in college than men; however, the proportion of women is the lowest in engineering, computer science, and physics disciplines at the undergraduate level. In 2014, only 19.8% of women enrolled in college were in engineering, and 14.5% of employed engineers were women in 2015 [11].

A study conducted by Gunderson et al. [12] shows how negative stereotyping of a female's math aptitude is passed to children by their parents and instructors. These clichés can prejudice young girls against math and hence, demolish their performance and enthusiasm for STEM subjects. There exist three main causes for the wide gender gap in enrollment in STEM courses [13]. These reasons include (a) a masculine culture that favors a sense of male inclusiveness in STEM, (b) inadequate experience for women in computer science, physics, and engineering, and (c) gender-based self-efficacy. The study also suggests that changing the cultural perception of women in STEM may benefit these fields and boost equality between the two genders by favoring the idea that both men and women can be successful.

Studies in the literature show that many women's social environment influences their educational orientation and the choices of their career paths, which tend to diverge away from STEM fields. Hence, researchers emphasize that the use of new teaching approaches could be beneficial in enhancing women's performance and self-confidence in STEM and changing the social perception of gender in the matter. A thorough literature review revealed that a few of the existing studies suggest the use of new technologies that facilitate the learning process. Particularly, technologies that involve immersive real-life class scenarios in industrial engineering. Furthermore, no existing study has paid particular attention to women's performance when implementing these technologies in education. Therefore, a gap subsists in conveying math, physics, and engineering concepts in a way that channels theory into practice. In the current study, a VR queuing theory teaching module is developed to demonstrate the efficacy and effectiveness of VR technology in boosting students learning. Fifty-six graduate and undergraduate students with no prior knowledge in queuing theory have participated in the study. The proposed VR module is then analyzed to investigate the impact of the VR environment on female engineers' in STEM.

The following section presents an overview of the literature concerning the use of VR in engineering education and applications, gender in STEM, the efficacy measures, and effectiveness measures of the proposed VR Module. Section three discusses the research design and methodology, along with the research question. The fourth section presents the results and analysis. Section five provides a discussion of the results of the study. Finally, section six presents the conclusions of the study and future work.

2. Related Work

This section is four-folds. First, the existing literature related to the benefits of VR in education and engineering applications is presented. Second, the literature pertaining to the women in STEM fields is reviewed. Then, a general overview of efficacy measures used in the study is discussed. Finally, an overview of the effectiveness measures utilized in the study is presented, along with the current challenges that need to be addressed. Additionally, the two gaps that subside in the literature concerning the current traditional teaching methods and their effect on women in STEM are also discussed in the final section.

2.1 Benefits of Virtual Reality in Education and Engineering Applications

Virtual Reality (VR) is an emerging technology and research area that carries the key to evolution in different fields such as business, education, engineering, and medicine. In education, the benefits of VR were thoroughly investigated. For instance, Bricken and Byrne [14] conducted a study to show the benefits of VR in bettering the learning process of students. With fifty-nine student participants, their results proved that VR enhances student's learning. The authors characterized VR as "a new way to use computers" and pointed that "VR eliminates the traditional separation between user and machine, providing more direct and intuitive interaction with information'' [14]. Based on the feedback from students, teachers, and researchers, their study outlined three educational areas where VR is applicable: Experimental education, constructivism, and social learning. Crosier et al. [15] also conducted a study involving fifty-one students from physics that favored VR to the existing teaching techniques in grasping radioactivity concepts. Foad and Whitman et al. [16, 17], on the other hand, used written and practical tests to assess students' perception towards conventional light microscopy (LM) and virtual microscopy (VM) using two random groups at the University of Tabuk in Saudi Arabia. The study revealed that students' performance with VM was better compared to LM and their grades were more uniform and less dispersed. The reason is that VR has advanced features that help students grasp difficult concepts better than any other traditional tool, which Mikropoulos and Natsis [3] revealed in their study. The authors stated that ''multisensory interaction channels, intuitive interactivity, and immersion'' are some of the main technological features in VR and they emphasized the idea that an 'Educational Virtual Environment' must be incorporated in the educational context. Another physical issue arises from the strict schedule of academic laboratories which affects the time flexibility and availability of materials for students. VR environment can resolve this issue and was proved to be the most convenient alternative for teaching subjects like robotics [18].

Many scholars have emphasized the importance of VR technological features [19–23]. These VR characteristics help students understand difficult concepts better than the two-dimensional representations through specialized peripherals such as space mouse, data gloves, and head trackers, offering the promising opportunity to learn in first-person, non-symbolic experience. Additionally, through their efforts, Crosier et al. [15] showed that high ability students were more capable of self-learning while lower ability students required more guidance and instructions. They concluded that VR-based learning modules could provide a suitable environment for better information grasping since it allows for:

(1) Data visualization and manipulation.
(2) Presentation of different perspectives.
(3) Risky situational sensing.
(4) Generation of three-dimensional (3D) concepts.

Engineering application is one of the critical fields that can benefit from extended realities (XR), especially VR, since most of the time, engineers are required to extend and analyze their findings in 2D settings through imagination and common sense. VR replaces this step by immersing its users with their findings in a virtual ambiance. Not only engineers are able to observe their findings from different angles in 3D settings, but they can also interact with them. Winn [19] cited multiples VR benefits allowing the direct interaction with virtual objects in ways not even possible in real-world settings: (1) Manipulation of size to increase/decrease the size of objects; (2) The ability to simulate ambiguous engineering concepts such us friction enabling engineers to make sense of it and to see the impact of it on other parameters instead of common sense interpretations; (3) Provide engineers with the first-person VR experience allowing them to stop feeling like in a simulation and start enjoying the benefits of VR through natural gestures like in real-world settings and even better. Furthermore, with the advancement of VR-related hardware and software, engineers can learn to build their own immersive environments and consequently build new customized features depending on the nature of their work. For instance, Goulding et al. [24] developed a VR tool that permits construction engineering students to build construction sequences in VR. The results suggest that engineering students handle construction tasks and gain experience faster when using VR than when using traditional tools. VR enables them to build full-scale buildings and the opportunity to critique their projects all in a virtual setting. This can help gain time, experience, and commit fewer errors (Low-cost) which are the main characteristics of success in the engineering fields.

In engineering education, experimental results from a study carried out by Mosterman et al. [25] to evaluate electrical engineering students showed that the time and assistance required for students who use virtual laboratories prior to physical laboratories are significantly less than students who use physical laboratories initially. The outcome of the study also indicated that the former group showed higher satisfaction with their laboratory experience than the latter group. For mechanical engineering students, Impelluso and Metoyer-Guidry [26] came up with a strategy called ''learner as instructional designer.'' This strategy engages students in developing VR-based models for traditional engineering concepts and allowed them to teach other students. The use of VR as a tool to explain engineering concepts to students increases motivation and profound conceptual understanding. Throughout their study, Ross and Aukstakalnis [27] confirmed that VR increases students' motivations while discussing the advantages of VR in engineering education by providing several existing applications (aircraft design, automotive design, architectural models, etc.) as well as some possible future applications.

Other than engineering education and applications, some other fields have also focused on VR applications to improve employee skills. For instance, to improve nursing educational skills, Smith and Hamilton [28] developed a VR simulation model as a supplement tool to increase students' awareness with regards to the critical steps in the nursing environment. Two groups had been formed as control and experimental. The latter group was allowed to use the VR program to practice catheterization skills. The descriptive findings of the study showed higher performance scores and perceived levels of preparedness in the experimental group than the control group. Along with William et al. [29], who mentioned that "virtual reality simulator (VRS) has the potential to bridge the gap between theory and practice for nursing students" and "increasing patient safety" and "reducing student anxiety?" Both studies encourage the implementation of VR to educate nurses and enhance their preparedness to face critical patient cases.

2.2 Women in STEM

Some of the challenges that lead to the under-representation of women in STEM field are: (1) *The Level of Self-Confidence* – Male students may have more confidence in studying and understanding engineering concepts than female students [12, 30], (2) *Interests and Level of Performance* – Low confidence levels could lower female students' interest in engineering fields [31, 32], (3) *Cultural Norms and Beliefs* – Female students may consider that engineering involves more physical effort and risk than other domains such as business [10, 33], and (4) *Concepts and Topics* – Engineering usually require more 'sequence-driven' concepts and topics and within a hierarchy of information. Furthermore, gender as a socio-cultural and psychological aspect of humans affect their creativity [34, 35], learning style [36], and the way they digest information [37].

Stadler et al. [38] conducted a study to examine the differences in learning physics concepts between male and female students and its impact on the learning outcome. The results showed that differences exist in the meaning of physics for males and females affected by (1) the relationship between daily life and the type of information; (2) body language; and (3) emotions. For this reason, and thanks to the advancement of human computer interface technologies, researchers started to develop new ways to account for these differences. For instance, Gunawan et al. [39] conducted a research involving three distinct high schools and used virtual laboratories to assess the impact of implementing these new technologies on the learning process of physics in a comparative study between males and females. Results showed that VR improves the creativity of both genders. Furthermore, the study also show that female students have greater verbal creativity than their male counterparts while male students outperform female students in figural creativity. Gunawan et al. [34] also conducted a quasi-experiment to assess the impact of virtual laboratories on the numerical, verbal, and figural creativity in learning physics concepts at four high schools with a total of 102 female and male students. This time, for figural and numerical creativity, female students outshined male students whistle both genders showed equal scores in the verbal creativity. Virtual laboratories can enhance creativity for both genders, which leads to believe that traditional teaching methods contribute to this under-representation as female students tend to dislike the structural rigidity of information that are presented [40]. With regard to the fourth challenge, *Concepts and Topics*, the current traditional teaching methods often fail to provide a practical understanding of engineering concepts, especially in industrial engineering curricula. In this study, we aim to investigate whether male and female students will learn differently from the VR module developed to teach queuing theory concepts.

2.3 Efficacy Measures of the Proposed Virtual Reality Module

The efficacy of a VR system determines its ability to perform the intended tasks in the immersive settings. Many studies have been conducted to measure the efficacy of VR systems through different evaluation approaches. For instance, Formosa et al. [41] designed a questionnaire-based evaluation method to assess the efficacy of a VR system in psychology education. The study assessed the user experience regarding a VR simulation using seven-point Likert-type questions over four sub-factors: fidelity, immersion, presence, and user buy-in. A thorough literature review revealed that most studies have not been capable of measuring the quality of VR systems in an adequate way. Furthermore, most of the assessment methods are limited to specific systems and lack user involvement in their evaluation, which appear to be major weaknesses in many studies. To overcome these issues, three more user-oriented questionnaires have been used in this study (simulation sickness questionnaire (SSQ) [42], system usability scale (SUS) [43], and presence questionnaire (PQ) [44]).

The simulation sickness questionnaire (SSQ) captures the user experience regarding the simulation sickness when exposed to a virtual environment. Kennedy et al. [42] have identified 21 sicknesses that can occur with VR simulations (Air force helicopter simulators, naval simulators, etc.) and were able to eliminate some symptoms with misleading indications and less frequent occurrences using factor analysis. The principal factors analysis with varimax rotation demonstrated that the three-factor solution is most appropriate for the reduced model that contains 16 symptoms. Three distinct clusters resulted from the analysis: nausea (general discomfort, salivation increasing, sweating, nausea, difficulty concentrating, stomach awareness, and burping), oculomotor disturbance (general discomfort, fatigue, headache, eye strain, difficulty focusing, difficulty concentrating, and blurred vision), and disorientation (difficulty focusing, nausea, fullness of the head, blurred vision, dizziness with eyes open, dizziness with eyes closed, and vertigo). Participants are given four levels (none, slight, moderate, and severe) to specify how much each symptom affects them during the VR study. These levels can be rated numerically between 0 and 4, with 0 being none and 4 being severe. The symptoms that are related to gastrointestinal distresses are presented by nausea, discomfort related to visual observations, and vestibular disturbances, are measured by the oculomotor and disorientation, respectively. The user score of each symptom is added together in each group and multiply with the relevant weight threshold (nausea: 9.54, oculomotor disturbance: 7.58, and disorientation: 13.92) to get the weighted sum. For the overall simulation sickness score, the sum of the scale scores is multiplied with 3.74. Table 1 represents the SSQ score and its corresponding categorization of the outcome of total scores.

The System Usability Scale (SUS) captures user feedback regarding how well the system is being set up. Brooke [43] has developed a tool that consists of ten items where the users can rate using a five-point Likert scale ranging from one (strongly disagree) to five (strongly agree). The ten statements in the SUS tool can be distinguished as positively worded or negatively worded when scoring the survey. Statements marked with odd numbers (1, 3, 5, 7, 9) specify the positive options while even numbers (2, 4, 6, 8, 10) postulate the negative options. For the positive options, the item score can be calculated by subtracting one from the scale position. For the negative options, the item score can be calculated by subtracting the scale position from five. In order to get the total SUS score, the summation of all item scores in both options needs to be multiplied by 2.5. Users are given the opportunity to evaluate the system usability immediately after the VR module, not allowing them to think for a long time. If the participants do not wish to answer a particular statement, they can choose the center point (3), and all items should be checked.

Table 1. SSQ score categorization

SSQ Score	Categorization
0	No symptoms
< 5	Negligible symptoms
5 – 10	Minimal symptoms
10 – 15	Significant symptoms
15 – 20	Symptoms are a concern
> 20	A problem simulator

User experience measures how well users engage in a simulated virtual environment. One of the widely used user experience tools is Presence Questionnaire (PQ), introduced by Witmer and Singer [44]. PQ consists of 22 questions that evaluate a user's experience with a VR system using a seven-point Likert scale ranging from zero (lowest level) to six (highest level). PQ contained five sub-scales of the user experience factors: involvement, immersion, visual fidelity, sound, and interface quality.

2.4 Effectiveness Measures of the Proposed Virtual Reality Module

The effectiveness of a VR module can be measured by the participants' performance in a virtual environment. The effectiveness of a VR simulation is also reflected by participants' knowledge gain and how much they feel motivated to use VR as a tool. To measure the effectiveness of the study, NASA TLX and intrinsic value subscale of the Motivated Strategies for Learning Questionnaire (MSLQ) are utilized.

Nasa Task Load Index is a multidimensional tool that measures the perceived workload of a task based on six aspects of performance [45]. The procedure is developed by the Human Performance Group at NASA Ames Research Center over a period of three years. Below are the six subscales that aid in calculating the overall workload score for a given task. *Mental Demand*: measures the mental effort required for a task (easy or demanding), *Physical Demand*: measures the physical effort required for a task (easy or demanding), *Temporal Demand*: measures the time pressure due to the pace of a task (slow or rapid), *Performance*: measures the accomplishment in achieving the goals of the task (good or poor), *Effort*: measure both mental and physical effort to deal with task (low or high), and *Frustration Level*: measures the user perception of dealing with a task (irritated or complacent). The overall NASA TLX score ranges between 0 and 100. Higher scores reflect a greater perceived workload for a given task.

The Motivated Strategies for Learning Questionnaire (MSLQ) is specifically designed to determine students' motivational orientation towards college courses [46]. To examine the students' degree of perception towards participating in a task, 14 intrinsic subscales were used with a seven-point (0~6) Likert scale.

Consequent to a thorough literature review, the two important gaps that need to be addressed to aid in enhancing the representation of women in STEM are as follows:

(1) The existing teaching approaches have not been successful in linking theoretical knowledge with practical implementation in an interactive and simulated way. This describes some of the cited challenges that female students confront.
(2) The present body of literature has not exhibited studies which utilize VR technologies in teaching the concepts of industrial engineering and manufacturing system and examines their effect on gender learning. Although students are able to learn the concepts and pass the exam, many of them will still not have a practical understanding or hands-on experience to apply many of the industrial engineering concepts such as queuing theory in real-life applications. This challenge negatively impacts female students since they are underrepresented in STEM and their fight to eliminate the cultural judgments and other genderism issues.

This set of challenges directs the focus and effort of this study toward minimizing the unfavorable impacts on female students. In an attempt to address the challenges along with the consideration of VR's benefits in education, this study investigates the impact of VR environments on the learning process on gender in STEM fields. To achieve the purpose of the study, a queuing theory VR teaching module was developed by Hamilton et al. [47] and was used to investigate the impact of the VR environment on males' and females' levels of knowledge gain in VR. The researchers also investigated the efficacy and other effectiveness measures of leveraging VR in teaching engineering concepts by measuring simulation sickness, system usability, user experience, NASA TLX, and level of post-motivation as a comparison between men and women.

3. Research Design and Methodology

Considering the VR benefits in education, discussed in the previous sections, this study integrates immersive VR to manufacturing systems concepts to enhance students' capabilities in dealing with virtual real-life complex manufacturing systems and their symptomatic problems. Specifically, this study aims to evaluate the efficacy and effectiveness of the VR module. This section contains a brief description of the VR queuing teaching module, research questions, and the design of the experiment.

3.1 The VR Queuing Theory Teaching Module

The study is based on immersive VR modules that demonstrate the queuing theory. Queuing theory is considered a well-known mathematical concept that deals with waiting lines (queues). It is a widely used topic in the domain of engineering to manage and improve systems' operations.

The VR module [47] is built using Unity game engine. Oculus Rift is used to connect the module with the immersive 3D environment. Before the module begins, the participants were asked to wear the headset and received instructions on the correct positioning. The headset lens was also adjusted along with adjusting the lens to ensure a clear view. The module began after the user selected the "Reach Here to Begin" button and the automatic audio recording explained the buttons and triggers on the touch controllers (see Fig. 1(a)). Once the user has successfully followed the commands on the hand controllers, he/she is prompted to move forward to click the "Reach Here to Continue" button to start the VR lecture. The VR lecture contained visual elements that demonstrated the difference between discrete and continuous data (see Fig. 1(b, c)) and three types of models in queuing theory (static, physical, and mathematical) (see Fig. 1(d, e, f)). Once the demonstration was over, the user was directed to interact with a glowing sphere to transport to a fast-food restaurant.

The fast-food restaurant scenario starts with an explanation of the theoretical concepts of queuing theory with visual aids. While the lecture is on, the user observes how queuing theory works in a restaurant with animated human models depending on the number of customers and servers available at a time (see Fig. 2). All the equations related to queuing theory appear on the space and move towards the relevant objects in the scene (see Fig. 2). During the VR lecture, the user may need some head and body movements to watch the lecture slides and equations while observing the animated customer lines in the restaurant (see Fig. 2). To lower the potential simulation sicknesses, the users were provided a pause/play option to receive a break during the lecture.

After the lecture and the demonstration are over, the students have a chance to manipulate the simulation by changing queuing parameters (inter-arrival time, arrival and service distributions, number of servers, etc.) to understand the queueing

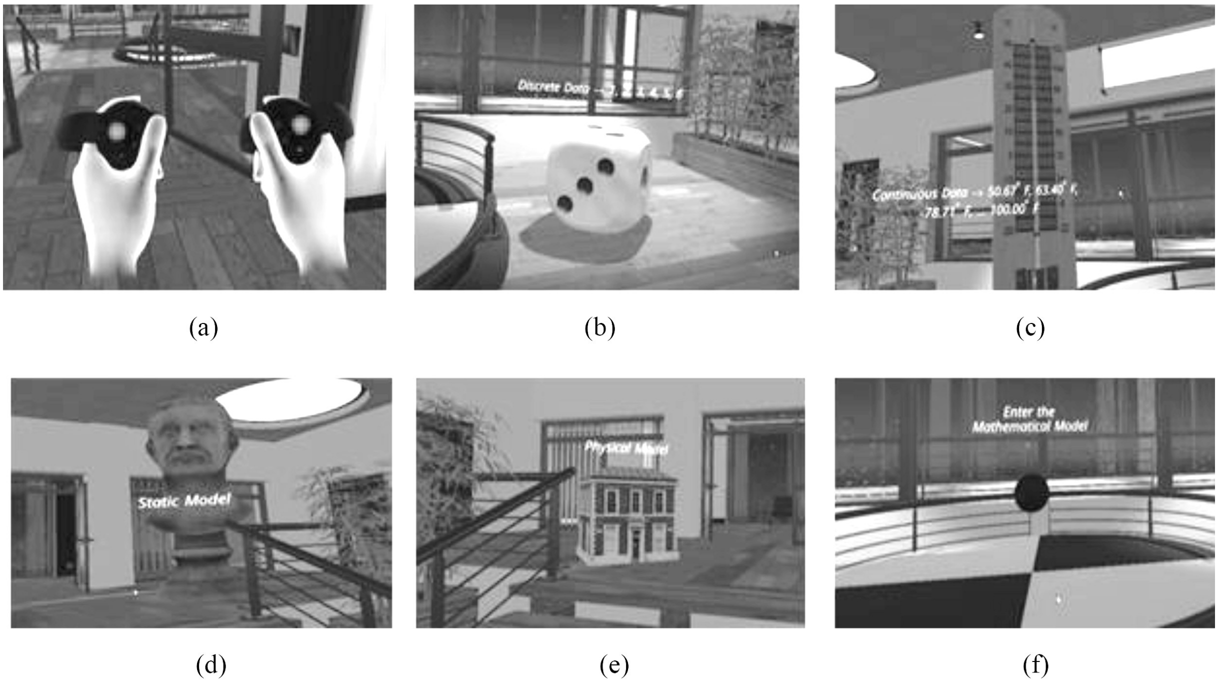
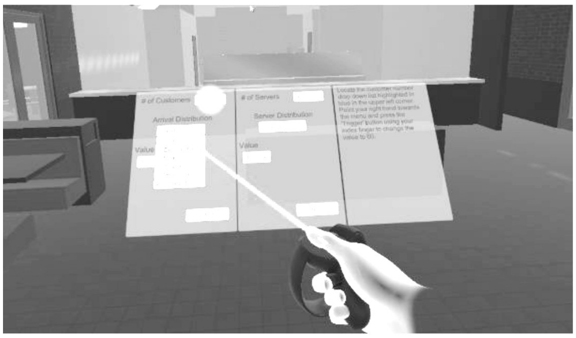


Fig. 1. Hand controllers and visual elements for queuing theory models.

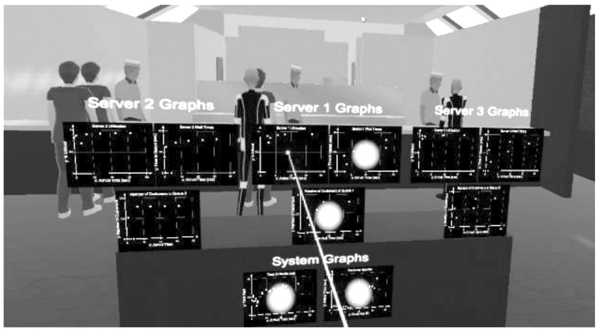
Fig. 2. Queuing theory lecture, animated equations, and preset human animations.

situations along with the graphical visualizations. To manipulate the parameters, the users are provided with a dashboard containing buttons along with both audio and text descriptions for menu options (see Fig. 3(a)). In addition to the customer queues, a set of graphs were provided to analyze real-time statistics of queuing theory model generated by users. The generated graphs interact with the controller tooltips to show point values (see Fig. 3(b)). For better comparison analysis, users can move the graphs to the center of the scene and group them together (see Fig. 3(b)). After the practical session ends, students must take a quiz that examine their knowledge of concepts in queuing theory and analytical skills. To start the quiz, a button is provided with the text "Reach Here to Begin" after students finished the visualization trials. The quiz consists of 14 conceptual questions and six analytical questions to test the students' knowledge gain. The students have the ability to use a virtual calculator and review the lecture if needed (see Fig. 3(c)). During the simulation, students are asked to fill the simulation sickness questionnaire three times during the experiment to monitor their health condition.

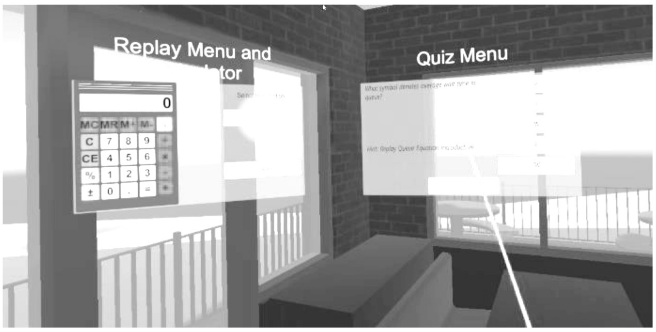
The current developed VR module is an extended version of the previous VR queuing theory developed by Ma et al. [7]. The extended VR version is more interactive and has improved user-friendly features for students to learn queueing theory. In the previous study, researchers have found three areas that need to be improved after evaluating the user experiences and comments to create a more user-friendly and effective learning environment for students to interact with the VR module. The new version of the VR module includes the below list of features that are developed based on the comments received from participants:



(a)



(b)



(c)

Fig. 3. Queuing dashboard, interactive graphs, and quiz interface.

- *Introduction and updated lecture*: In the new version of the VR module, a tutorial is provided on how to use the controllers followed by a detailed lecture that covers the following topics: the concept of queuing theory, explanation of discrete data vs. continuous data, differences between static models, and physical discrete event models and mathematical models that represent queuing theory.
- *Simulation*: The lecture audio is also added along with preset values of the queuing engine to generate a queue during the lecture. Pause/Play button for breaks during the lecture, and audio to explain menu options for changing values to the queuing engine are added to the new version. Graphs are changed from being static to being interactive with a tooltip to show point values.
- *Practice/Quizzes*: In the updated version of the model, users can take multiple quizzes in the VR module to test their knowledge in queuing theory.

3.2 Research Questions

To meet the study objectives, the following research questions are formulated:

3.2.1 Efficacy Research Questions

Are there any significant differences between male and female students participating in the new version of the VR teaching module regarding efficacy measures, including, (1) simulation sickness, (2) system usability, and (3) user experience?

3.2.2 Effectiveness Research Questions

Are there any significant differences between male and female students participating in the new version of the VR teaching module regarding effectiveness measures, including (1) quiz score, (2) NASA TLX performance, and (3) level of post-motivation?

3.3 Design of Experiment and Study Sample

A total of 56 students participated in the study, including 21 females, 32 males, and three participants declined to specify their gender. Seventy-five percent (n = 42) of the participants were graduate students. Thirty-nine percent (n = 14) were majoring in industrial engineering. The majority (n = 38, 67%) of the participants were international students. The data collection process took approximately one day to finish.

Each participant was required to complete ten background questions, including five demographical questions (academic classification, major, gender, race, and origin) and five knowledge-based questions (statistical background, system simulation knowledge, restaurant management

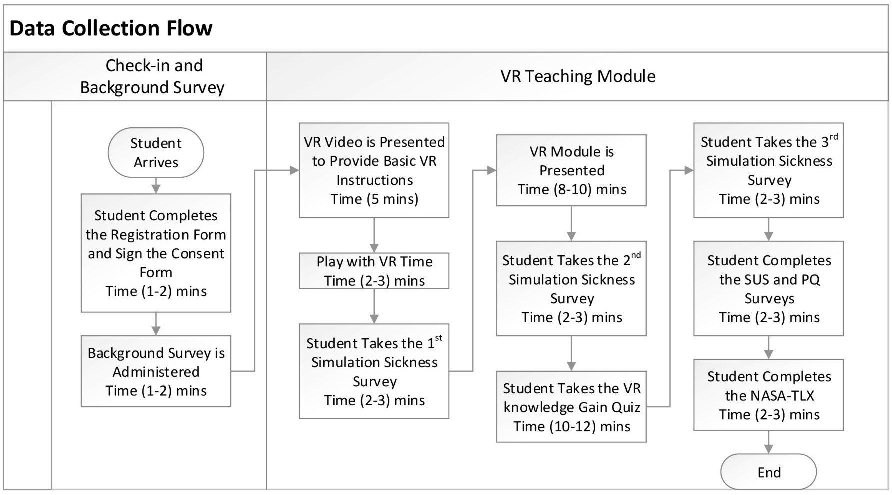


Fig. 4. Data collection process flow of the study.

knowledge, VR experience, and video game playing experience). The survey used a five-point Likert scale (0–4), which represented their knowledge as 'none' (0), 'basic' (1), 'average' (2), 'above average' (3), or 'expert' (4). Thirty-six of the participants (64%) had below-average prior experience in VR; thirty-eight (69.64%) had below-average prior experience in restaurant management/service experience. All participants had an equal opportunity to take part in the study, and they were selected upon the time of their arrival.

Fig. 4 shows the process flow of the data collection in the experimental study. The study aims to investigate the differences between males and females regarding the VR module efficacy (SSQ, SUS, and PQ) and to measure the VR effectiveness between males and females through knowledge gain quiz, NASA TLX, and post-motivation survey.

Upon arrival, the students completed the registration forms and signed the consent forms. Then, they were asked to take a survey to evaluate the socio-demographic background of the participants. Next, the participants were exposed to the VR environment to become familiar with its equipment before the study began. Upon completion of the training part, the students were asked to take the first simulation sickness survey. The purpose of the SSQ questionnaire is to assess the students' comfort level while engaging in the VR world. From this, the VR lecture module was provided to the participants allowing them to learn the concepts of the queueing theory through the VR module. When they completed the VR module learning, they were asked to take the second SSQ questionnaire followed by a quiz. The exercise aimed to evaluate the participants' knowledge gain from the VR lecture. In this study, knowledge gain is measured through conceptual quiz score, the analytical quiz score, and overall quiz score. The quiz consisted of 20 questions, which tested the knowledge of single and multi-queuing theory. Among the quiz questions, 14 were conceptual, and six were designed to test the analytical skill of a special case of batch queuing. The scores were calculated based on the correct answers for each quiz question.

Upon completion of the quiz, the students were given the last SSQ questionnaire to record the health data with respect to the VR environment interaction. After, the students took the survey related to the system usability and the user experience. The system usability survey measured participants' perception of the VR module. Ten system usability questions with a five-point Likert scale were used to collect the data. The user experience survey measured the participants' experience of using the VR teaching module and 22 questions with a seven-point Likert scale was used to collect the data. The post-motivation survey was utilized to assess the students' level of motivation after going through the VR experience. Fourteen questions with a seven-point Likert scale were used to collect the responses. In the end, students were asked to take the NASA TLX online quiz to assess the students' perceived value towards the task load.

4. Results and Analysis

The results of the efficacy and effectiveness measures of the proposed VR module is discussed in this section. The results were compared between males and females in all measures to explore the gender differences.

4.1 Efficacy Results of the Virtual Reality Module

Three survey/questionnaire instruments were used to assess the efficacy of the VR module, namely, simulation sickness, system usability, and user experience surveys. For an individual, switching from real-world to the three-dimensional virtual world, cause a mismatch in sensory inputs, which leads to simulation sickness. These symptoms vary from slight to severe and can cause fatigue and severe vomiting [48]. The efficacy measures are used to capture the level of discomfort the VR module may cause. The tools used for data collection are categorized questionnaires, and the output scores are represented and analyzed.

4.1.1 Simulation Sickness

Simulation sickness is the discomfort experienced by users in a simulated environment. To ensure the safety of the developed simulator, the SSQ questionnaires are distributed in three stages during the experiment: at the beginning, in the middle, and at the end of the experiment. The questionnaire helps to collect information regarding sixteen possible user symptoms, the degree of discomfort that participants might experience in the virtual environment. Each response scores for symptoms are none, slight, moderate, or severe, indicated by 0, 1, 2, and 3, respectively. These sixteen symptoms are divided into three groups, such that each group represents a unique indicator of simulation sickness. These groups are nausea, oculomotor discomfort, and disorientation. As described in section 2.3, to get the weighted sum of scores, user scores for each subgroup are summed and multiplied with the weights of 9.54, 7.58, and 13.92, respectively. The total simulation sickness score is multiplied with the assigned weight of 3.74 to get the overall SSQ score. SSQ scores can be considered as an aggregate analysis score for all participants. This indicates the severity of symptoms and helps fix troublesome indicators. Table 2 provides calculated SSQ scores based on all students' responses.

Table 2 indicates that overall, the VR module can cause simulation sickness at the beginning of the study, during the study, at the end of the study, with a steady increment in sickness for male participants. However, female participant reported fewer problematic symptoms at the beginning of the study with an increase in the severity of the symptoms with time. To determine the differences in both genders regarding simulation sickness, the independent-samples t-test is conducted for all the simulation sickness indicators and the total scores are calculated to identify the average difference in the mean scores for male and female participants. In the first SSQ, participants' scores indicate no significant difference in all groups even though females reported fewer symptoms than males. Females reported higher symptoms in the second and third SSQs compared to their counterparts. However, the p value confirmed that there is no statistically significant difference in the scores between male and female students in both SSQs.

4.1.2 System Usability

To specify the usability of a system, it is essential to know the possible users, the tasks that users perform, and the characteristics of different environments that the system deals with. ISO 9241-11 [49] recommends that three measures of effectiveness, efficacy, and satisfaction, should be capable of assessing the usability of a system. SUS questionnaire is considered as a useful tool that can capture all usability components. A total SUS score lower than 68 indicates that the user satisfaction is below-average while a total SUS score higher than 68 suggests an above-average user satisfaction. In this study, the total average SUS score for both male (M = 76.00) and female (M = 77.29) groups suggest above-average user satisfaction. In addition, the total average usability score of 76.47 (SD = 12.26) also indicates that participants perceived above-average user satisfaction from the VR queuing theory module. Table 3 presents the mean scores and standard deviations of the user responses for the ten scale items in the SUS. The first seven items of the SUS questionnaire are positively worded, and the last three items are negatively worded. The average score of each item shows that both males and females agreed (>3) with the positive description of the proposed module and disagreed (>3) with the negatively described items. To investigate the impact of gender on system usability, an independent-samples t-test was carried out. The results confirmed that the usability scores of female students are not significantly different from the usability score of their counterparts in all scale items. Similarly, the p value recommended that male and female students are not significantly

Table 2. Simulation sickness questionnaires scores

	SSQ 1		t-test		SSQ 2		t-test		SSQ 3		t-test	
	M	**Fe**	**t**	***p***	**M**	**Fe**	**t**	***p***	**M**	***Fe***	***t***	***p***
Nausea	14.3	11.6	0.49	0.63	14.7	21.8	1.43	0.16	21.9	28.6	0.87	0.39
Oculomotor	19.0	16.8	0.36	0.72	20.5	22.7	0.36	0.72	28.1	38.4	1.22	0.23
Disorientation	31.3	18.9	1.40	0.17	31.3	27.8	0.37	0.72	41.2	46.7	0.39	0.70
Total	23.4	17.9	0.83	0.41	24.3	27.3	0.45	0.65	33.5	42.7	0.93	0.36

Table 3. System usability scale scores

System Usability Scale Items	Male		Female	
	Avg. Score (out of 5)	SD	Avg. Score (out of 5)	SD
1. I would imagine that most people would learn to use this VR module very quickly.	3.92	1.18	4.14	0.66
2. I found the various functions (e.g., sound, pictures, control) in this VR module were well integrated.	3.83	0.87	4.21	0.97
3. I felt very confident using this VR module.	4.08	0.88	3.64	1.01
4. The VR modules helped me to establish the linkage between the concept of queueing theory and practice.	3.92	0.78	4.00	0.68
5. I found this VR module was easy to use.	3.92	0.88	3.79	0.97
6. I would like to use VR in other courses.	3.83	1.13	3.86	1.17
7. I think that I would like to use this VR module to learn the queueing theory.	3.79	0.98	3.57	1.09
8. I think I would need the support of a technical people to use this VR module.	3.83*	1.09	3.79*	1.12
9. I should learn more VR base knowledge before I use the VR module.	3.25*	1.15	3.64*	1.01
10. I thought there was too much inconsistency in this VR module.	3.63*	0.97	4.00*	1.24
Total System Usability	3.80	0.67	3.86	0.53

* Reverse coded to be consistent with other questions.

different from each other regarding the total SUS score (t = 0.50, p = 0.62).

4.1.3 User Experience

User experience is used to measure the perception of involvement and immersion in a virtual environment. This study utilizes PQ to evaluate user experience. PQ involves five sub-scales: involvement, immersion, visual fidelity, interface quality, and sound captured through 22 questions using a scale between 0–6. Table 4 displays the gender-wise average scores for each of the sub-scales and the average for all items in PQ. Except for interface quality, all subscales indicate above-average (>3) user experience in the male students while female students experienced above-average (>3) user experience in all subscales. The average score for all subscales given by male participants is 4.15, and the response given by female participants have an average score of 4.50. The independent-samples t-test was carried out to investigate the differences between male and female students regarding the user experience. The independent-samples t-test results indicated that females are not significantly different from male students regarding the five subscales of user experience in the VR module. However, for all 22 PQ items, the t-test results confirmed that the female students are significantly different from their counterparts with a better user experience (t = 2.35, p = 0.02).

4.2 Effectiveness Results of the Virtual Reality Module

The effectiveness of the VR module for teaching queueing theory concept was evaluated using three methods, including students' knowledge gain quiz, NASA TLX work-load assessment, and post-motivation survey.

4.2.1 Knowledge Gain

The knowledge gain section consists of 20 multiple-choice questions to assess the students' conceptual and analytical skills regarding the queueing theory. Fourteen conceptual questions and six analytical questions are designed to measure the knowledge gain. Table 5 shows the descriptive statistics of the knowledge gain quiz. The results indicate that overall, students performed better on the conceptual part than they did on the analytical portion of the quiz. Moreover, female students outperformed male students in the analytical, conceptual, and overall scores. However, the quiz scores between male and female students are not significantly

Table 4. Presence questionnaire scores

Subscales of PQ	Items	Male		Female		t-test	
		Avg. Score (out of 6)	SD	Avg. Score (out of 6)	SD	t	*p*
Involvement	1, 2, 3, 4, 5, 6, 7, 10, 13	4.57	0.71	4.91	0.59	1.51	0.14
Immersion	8, 9, 14, 15, 16, 19	4.10	0.51	4.32	0.48	1.33	0.19
Visual Fidelity	11, 12	4.67	0.87	5.07	0.76	1.45	0.16
Interface Quality	17, 18	2.81	1.41	3.39	0.90	1.38	0.18
Sound	20, 21, 22	4.58	0.73	4.79	1.08	0.69	0.49
Total	1-22	4.15	0.43	4.50	0.47	2.35	0.02

Table 5. Quiz score statistics

	Male		Female		t-test	
	Avg. Score	**SD**	**Avg. Score**	**SD**	**t**	***p***
Conceptual Score	73%	21%	78%	18%	0.73	0.47
Analytical Score	52%	23%	55%	32%	0.35	0.73
Overall Score	68%	17%	73%	17%	0.79	0.44

different in the conceptual portion of the quiz (t = 0.73, p = 0.47), the analytical portion of the quiz (t = 0.35, p = 0.73), and overall quiz scores (t = 0.79, p = 0.44). Overall, both the male and female students gained a high percentage of conceptual knowledge and a fair percentage of analytic knowledge regarding queueing theory by the VR module.

For a better overview of how students performed in the quiz, Table 6 was created with the correct answers and corresponding percentages for the male and female students. The results showed that the male students performed better in questions 1, 5, 7, 11, 14, 17 than female students. On the other hand, a higher percentage of female students correctly answered the remaining questions than their counterparts. Additionally, the difference between male and female students regarding the correct answer percentage ratio for questions 1, 2, 4, 8, 12, 13, 15, 16–20 are noticeable. For example, in question 1, 60 percent of male students correctly answered this question, while only 41 percent of female students answered accurately. Moreover, male students perform well (more than 70% accuracy) on questions 3, 5–7, 9–17, while they performed poorly on questions 18-20 (less than 50% accuracy). Female students perform well on questions 2–4, 6–9, 11–17, whereas they performed poorly on the first question.

Fig. 5 depicts the radar plot that demonstrates the total score of the conceptual and analytical portions in the queueing theory quiz on a 0.0–100.0% scale. It is visible that most of male students performed better in the conceptual quiz than the analytical quiz. On the other hand, the radar plot for female inferred that they scored higher in the conceptual portion of the quiz than in the analytical part.

4.2.2 NASA TLX Assessment

The overall NASA TLX index score of female students (M = 61.45, SD = 14.68) was found to be higher than male students (M = 49.15, SD = 22.92) in the study, concluding that the female group perceived on average 61% work effort to respond to the quiz questions while male group perceived on average 49% work effort to answer the quiz questions (see Fig. 6). Table 7 displays the weighted scores of six aspects of performance, including effort, frustration, mental demand, performance,

Table 6. Question-wise correct answers for knowledge gain quiz

		Male		Female	
	Question	**No. of Correct Answers**	**%**	**No. of Correct Answers**	**%**
Conceptual	1	18	0.60	7	0.41
	2	15	0.50	12	0.71
	3	21	0.70	12	0.71
	4	16	0.53	13	0.76
	5	21	0.70	11	0.65
	6	23	0.77	14	0.82
	7	27	0.90	14	0.82
	8	20	0.67	13	0.76
	9	25	0.83	15	0.88
	11	25	0.83	14	0.82
	12	25	0.83	16	0.94
	13	23	0.77	16	0.94
	14	28	0.93	15	0.88
	15	23	0.77	15	0.88
Analytical	10	15	0.50	10	0.59
	16	25	0.83	16	0.94
	17	26	0.87	12	0.71
	18	11	0.37	9	0.53
	19	14	0.47	10	0.59
	20	12	0.40	9	0.53

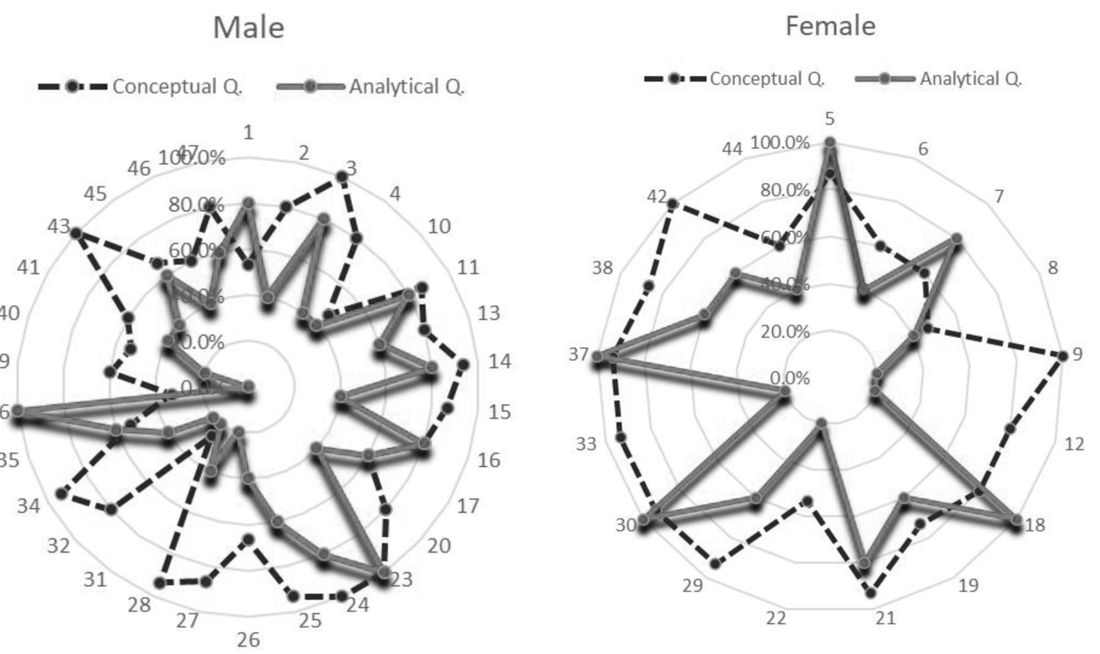


Fig. 5. Gender-wise knowledge gain quiz answers.

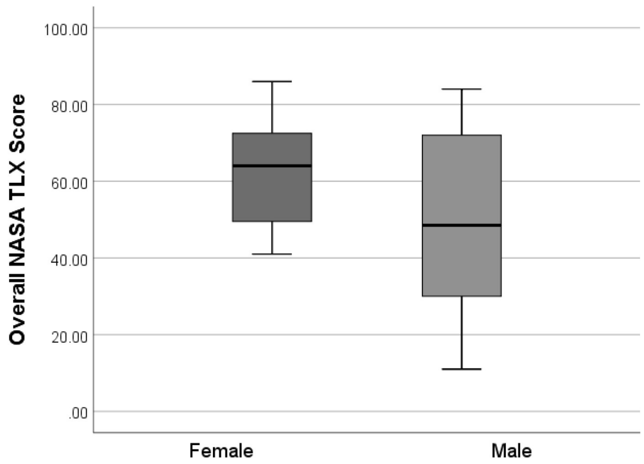


Fig. 6. Gender-wise overall NASA TLX performance.

Table 7. NASA TLX aspects of performance

	Male		**Female**		**t-test**	
NASA TLX Subscales	**Avg. Score**	**SD**	**Avg. Score**	**SD**	**t**	***p***
Physical Demand	0.05	0.07	0.08	0.09	1.03	0.31
Temporal demand	0.15	0.10	0.12	0.08	0.73	0.47
Mental Demand	0.25	0.07	0.22	0.14	0.77	0.45
Frustration	0.11	0.09	0.13	0.12	0.42	0.68
Effort	0.20	0.07	0.16	0.10	1.19	0.24
Performance	0.22	0.10	0.25	0.06	0.81	0.43
Overall NASA TLX Score	49.15	22.92	61.45	14.68	1.60	0.12

physical demand, and temporal demand in male and female groups. The independent-samples t-test was carried out on the overall score and the results indicate that females are not significantly different from males regarding the overall score ($t = 1.60$, $p = 0.12$). Furthermore, the t-test results confirmed that women perceived the same work effort as men across all different aspects of NASA TLX sub-

Table 8. Post-motivation scores

	Male		Female	
Post-motivation Scale Items	**Avg. Score (out of 6)**	**SD**	**Avg. Score (out of 6)**	**SD**
1. Compared with other students in this lecture I expect to do well in the knowledge practice.	4.33	0.96	3.86	1.75
2. I'm certain I can understand the ideas taught in this VR module.	4.83	0.96	4.29	1.59
3. I expect to do very well in future similar VR modules.	4.67	1.05	4.64	1.22
4. I am sure I can do an excellent job on the problems and tasks assigned in future VR modules like this.	4.54	0.98	4.43	1.45
5. I think I will receive a good score in future VR modules like this.	4.50	0.93	4.57	1.22
6. Compared with other students in this group, I think I know a great deal about the queueing theory.	4.13	1.26	3.36	2.02
7. I know that I will be able to learn the material better in future VR modules.	4.33	1.13	4.43	1.34
8. I prefer tasks that are challenging so I can learn new things after taking VR modules like this.	4.33	1.27	4.36	1.22
9. It is important for me to learn what is being taught in this VR module.	4.42	1.32	4.29	1.27
10. I like what I am learning in this VR module.	4.71	1.08	4.36	1.69
11. I think I will be able to use what I learn in this VR module in other classes.	4.29	1.68	4.29	1.68
12. I think that what I am learning in this VR module is useful for me to know.	4.79	1.32	4.50	1.74
13. I think that what we are learning in this VR module is interesting.	4.63	1.10	4.64	1.22
14. Understanding queueing theory from this VR module is important to me.	3.96	1.49	3.62	2.06
Total Motivation	4.46	0.86	4.27	1.16

scales. The results display that, on average, mental demand and performance contributed most to the overall score for women and men. In both groups, on average, physical demand was found to be the lowest contribution for overall index scores in regard to performing the quiz task.

4.2.3 Level of Motivation

The post-motivation survey used to measure the students' motivation after going through the VR experience. Fourteen questions with a seven-point Likert scale (coded 0 to 6) were utilized to collect the data from male and female participants. Male students scored higher than female students in questions 1–4, 6, 9, 10, 12, and 14, while female students scored higher than male students in remaining questions. The independent-samples t-test results of these 14 questions showed that the level of post-motivation of both groups is not significantly different from each other. With respect to total post-motivation of students, males scored higher than females; however, the difference between them is not significant ($t = 0.34$, $p = 0.56$). The total post-motivation scores of male and female students were 4.46 and 4.27 (out of 6), respectively, which means the VR module created positive motivation in learning the queueing theory for both male and female students.

5. Discussion

Educational VR modules as a classroom teaching method, help students to be innovative; it provides the student with an opportunity to not only to feel the first-person VR experience but to experience the impact of parameter change rather than constructing the rational explanation. The finding of this study suggests that both genders perform well in the presence of VR environments, which is consistent with the outcome of the previous studies [3, 8, 14, 15, 16, 17]. These findings validate the aptitude of VR laboratories in generating virtual learning environments that enhance the learning process of students and facilitate their understanding of the engineering concepts that are often difficult to grasp. Therefore, the VR laboratories can benefit female students in learning engineering concepts, particularly queuing theory concepts. Female students performed better than their male counterparts in learning queuing theory concepts using the module developed in the study, as shown in the results section. The result of our study regarding academic motivation in the sense that VR module can increase the students' motivations in understanding and practicing of engineering concepts, is consistent with other studies [25–27] that showed VR module enhance engineering students' motivation and satisfaction. Furthermore, the ability of VR to create an ample virtual space with advanced visualization features enhance students' perception of the topics taught to them [2, 4–6, 19–23]. In our study, both male and female students experienced well in different aspects of VR environment such as involvement, immersion, visual fidelity, and sound in addition to system usability scale (e.g., how well the system is being set up).

In order to abolish the negative stereotypes on female's competences in STEM fields, a study conducted by Gunderson et al. [12] suggests that changing the cultural perception of women in STEM may have a positive impact in both female's performance in STEM and the social perception of females in the matter. The current study advocates the idea that both men and women can be equally successful in these fields. The quiz scores of our VR queuing theory module indicated that female students outperformed male students in the overall scores of the conceptual and the analytical parts.

A thorough investigation has revealed that the design of the VR module had an impact on the student scores and simulation sickness. Many students easily answered the conceptual questions as they can understand the concept of queueing theory better with powerful visualization of the VR module. On the other hand, analytical questions require calculations and can relatively be time-consuming, compared to the conceptual questions, causing the students to perform poorly. Majority of the students mentioned that the animated equations which they replay while working on analytical quiz increases simulation sickness. Additionally, the low scores on analytical questions may indicate that the VR module should be improved to prepare the students more for analytical type questions. One idea can be designing more practice questions for students before taking the quiz.

Although this study presents insightful findings of the potential of using VR in classroom settings, it has several limitations. First, the smaller sample size. Future studies should take into consideration the recruitment of more students to acquire more statistical sound conclusions. Second, the study did not include a control group to which the findings of the VR study are to be contrasted with the traditional teaching method. The future versions should include a control group that will not be subjected to VR testing. The third, most of the students come from an engineering background with the same concentrations. This might create a bias in the study results. Future in-depth research on the subject is essential to validate these findings through repeated experiments that include far more participants with varying backgrounds and with the presence of control groups. More emphasis should be on qualitative research, including focus groups and semi-structured interviews, to gain a better insight into the learning process of students. Finally, more measures need to be implemented in future work like Net Promoter Score (NPS) and cheaper VR technologies alternatives.

6. Conclusion and Future Work

This study investigates how male and female students respond to learning engineering concepts in a more sophisticated environment. For this reason, a VR module has been developed to enhance the teaching process of manufacturing system concepts. The idea of using VR technologies was to create an immersive real-life experience potentially and to engage students in learning engineering concepts actively. The results were analyzed to investigate the impact of efficacy and effectiveness on both genders. The efficacy and effectiveness of the VR module determine the ability to perform well in an immersive environments. Based on the past studies, five evaluation methods are used along with a quiz to assess of the module.

The efficacy measures indicate that the students are satisfied in terms of the systems usability and the user experience. However, the researchers received some negative feedback from the students regarding the design aspect of the VR module, causing some students to experience a high level of simulation sickness. Both males and females show no significant difference in regrading simulation sickness and usability of the system while females recorded significantly higher user experience than their male counterparts. The effectiveness measures suggested that the developed VR module created positive motivation in learning queuing theory for both males and females. Furthermore, in the same way, both male and female students scored well in the conceptual quiz than the analytical quiz. The results confirmed that both male and female students have no significant difference in all three effectiveness measures.

Future studies will focus on more economical and powerful VR simulations that can provide better real-life experience in an interactive environment. For the current study, the researchers used a single device type (Oculus Rift) to link with immersive VR environment. Other devices, i.e., HTC Vive, and other cheaper alternatives such as Google Cardboard could be explored in future studies. Furthermore, the researchers used the Unity game engine to create the VR teaching module because of its capability in providing an interactive VR environment. Other software packages like Simio® or FlexSim® will be used to create interactive VR environments in future studies. New features (Virtual whiteboard and markers, path visualizers, dashboard, etc.) will be added to enhance the user experience and lower the simulation related sicknesses. Mainly, the researchers will concentrate on embedding new mitigation techniques to future VR studies to reduce sicknesses. Researchers believe such VR modules can better prepare "work-ready" students since VR helps to establish the much-needed linkage between theory and practice and gives students an effective virtual opportunity to validate the theory themselves.

References

1. M. A. Gigante, Virtual reality: Definitions, history and applications. In: R. A. Earnshaw, M. A. Gigante, H. Jones (Eds), *Virtual reality systems*, Academic Press, London, pp. 3–15, 1993.
2. V. S. Pantelidis, Virtual reality and engineering education, *Computer Applications in Engineering Education*, **5**(1), pp. 3–12, 1997.
3. T. A. Mikropoulos and A. Natsis, Educational virtual environments: A ten-year review of empirical research (1999–2009), *Computers & Education*, **56**(3), pp. 769–780, 2011.
4. V. Ramasundaram, S. Grunwald, A. Mangeot, N. B. Comerford and C. M. Bliss, Development of an environmental virtual field laboratory, *Computers & Education*, **45**(1), pp. 21-34, 2005.
5. O. Çaliskan, Virtual field trips in education of earth and environmental sciences, *Procedia–Social and Behavioral Sciences*, **15**, pp. 3239–3243, 2011.
6. P. N. A. Barata, M. R. Filho and M. V. A. Nunes, Consolidating learning in power systems: virtual reality applied to the study of the operation of electric power transformers, *IEEE Transactions on Education*, **58**(4), pp. 255–261, 2015.
7. J. Ma, R. Jaradat, O. Ashour, M. Hamilton, P. Jones and V. L. Dayarathna, Efficacy Investigation of Virtual Reality Teaching Module in Manufacturing System Design Course, *Journal of Mechanical Design*, **141**(1), pp. 1–13, 2019.
8. Daily Mail, https://www.dailymail.co.uk/femail/article-2051095/As-brand-suggests-girls-bad-math–just-fashion-let-idea.html, Accessed 18 March 2018.
9. E. Reuben, S. Paola and Z. Luigi, How stereotypes impair women's careers in science, *Proceedings of the National Academy of Sciences*, **111**(12), pp. 4403–4408, 2014.
10. C. A. Moss-Racusin, J. F. Dovidio, V. L. Brescoll, M. J. Graham and J. Handelsman, Science faculty's subtle gender biases favor male students, *Proceedings of the National Academy of Sciences*, **109**(41), pp. 16474–16479, 2012.
11. National Science Foundation (NSF), https://www.nsf.gov/statistics/2017/nsf17310/digest/about-this-report/, Accessed 18 March, 2018.
12. E. A. Gunderson, G. Ramirez, S. C. Levine and S. L. Beilock, The role of parents and teachers in the development of gender-related math attitudes, *Sex roles*, **66**(3–4), pp. 153–166, 2012.
13. S. Cheryan, S. A. Ziegler, A. K. Montoya, L. Jiang, Why Are Some STEM Fields More Gender Balanced Than Others?, *Psychological Bulletin*, **143**(1), pp. 1–35, 2017.
14. M. Bricken and C. M. Byrne, Summer students in virtual reality: A pilot study on educational applications of virtual reality technology, In A. Wexelblat (Ed.), *Virtual reality: applications and explorations*, Academic Press, Boston, MA, pp. 199–218, 1993.
15. J. K. Crosier, S. V. G. Cobb, J. R. Wilson, Experimental comparison of virtual reality with traditional teaching methods for teaching radioactivity, *Education and Information Technologies*, **5**(4), pp. 329–343, 2000.
16. A. F. Foad, Comparing the use of virtual and conventional light microscopy in practical sessions: Virtual reality in Tabuk University, *Journal of Taibah University Medical Sciences*, **12**(2), pp. 183–186, 2017.
17. L. Whitman, D. Malzahn, V. Madhavan, G. Weheba and K. Krishnan, Virtual reality case study throughout the curriculum to address competency gaps, *International Journal of Engineering Education*, **20**(5), 690–702, 2004.
18. F. A. Candelas, S. T. Puente, F. Torres, F. G. Ortiz, P. Gil and J. Pomares, A virtual laboratory for teaching robotics, *International Journal of Engineering Education (Special Issue on Distance Controlled Laboratories and Learning Systems)*, **19**(3), pp. 363–370, 2003.
19. W. D. Winn, A conceptual basis for educational applications of virtual reality, Technical Publication R-93-9, Human Interface Technology Laboratory of the Washington Technology Center, University of Washington, Seattle, 1993.
20. M. D. Dickey, Brave new (interactive) worlds: A review of the design affordances and constraints of two 3D virtual worlds as interactive learning environments, *Interactive Learning Environments*, **13**(1–2), pp. 121–137, 2005.
21. C. Dede, M. C. Salzman, R. B. Loftin and D. Sprague, Multisensory immersion as a modeling environment for learning complex scientific concepts, *Modeling and simulation in science and mathematics education*, Springer, New York, NY, pp. 282–319, 1999.
22. H. R. Neale, D. J. Brown, S. V. G. Cobb and J. R. Wilson, Structured evaluation of virtual environments for special-needs education, *Presence: teleoperators and virtual environments*, **8**(3), pp. 264–282, 1999.
23. W. B. Lee, J. G. Li and C. F. Cheung, Development of a virtual training workshop in ultra-precision machining, *International Journal of Engineering Education*, **18**(5), pp. 584–596, 2002.
24. J. Goulding, W. Nadim, P. Petridis and M. Alshawi, Construction industry offsite production: A virtual reality interactive training environment prototype, *Advanced Engineering Informatics*, **26**(1), pp. 103–116, 2012.
25. P. J. Mosterman, M. A. Dorlandt, J. O. Campbell, C. Burow, R. Bouw, A. J. Brodersen and J. R. Bourne, Virtual Engineering Laboratories: Design and Experiments, *Journal of Engineering Education*, **83**(3), pp. 279–285, 1994.
26. T. Impelluso and T. Metoyer-Guidry, Virtual reality and learning by design: Tools for integrating mechanical engineering concepts, *Journal of Engineering Education*, **90**(4), pp. 527-534, 2001.
27. W. A. Ross and S. Aukstakalnis, Virtual Reality: Potential for Research in Engineering Education, *Journal of Engineering Education*, **83**(3), pp. 287–291, 1994.
28. P. C. Smith and B. K. Hamilton, The Effects of Virtual Reality Simulation as a Teaching Strategy for Skills Preparation in Nursing Students, *Clinical Simulation in Nursing*, **11**(1), pp. 52–58, 2015.
29. A. William, V. L. Vidal and P. John, Traditional Instruction versus Virtual Reality Simulation: A Comparative Study of Phlebotomy Training among Nursing Students in Kuwait, *Journal of Education and Practice*, **7**(9), pp. 18–25, 2016.
30. D. N. Beede, T. A. Julian, D. Langdon, G. McKittrick, B. Khan and M. E. Doms, Women in STEM: A gender gap to innovation, *Economics and Statistics Administration Issue Brief*, pp. 04–11, 2011.
31. E. Pronin, C. M. Steele and L. Ross, Identity bifurcation in response to stereotype threat: Women and mathematics, *Journal of Experimental Social Psychology*, **40**(2), pp. 152–168, 2004.

32. J. L. Glass, S. Sassler, Y. Levitte and K. M. Michelmore, What's so special about STEM? A comparison of women's retention in STEM and professional occupations, *Social forces*, **92**(2), pp. 723–756, 2013.
33. J. W. Wimer, C. S. Ridenour, K. Thomas and A. W. Place, Higher order teacher questioning of boys and girls in elementary mathematics classrooms, *The Journal of Educational Research*, **95**(2), pp. 84–92, 2001.
34. G. Gunawan, N. M. Y. Suranti, N. Nisrina, L. Herayanti and R. Rahmatiah, The effect of virtual lab and gender toward students' creativity of physics in senior high school, *Journal of Physics: Conference Series*, **1108**(1), p. 012043, 2018.
35. J. W. Santrock, Psikologi Pendidikan, Jakarta: Kencana Prenada Media Group, 2008.
36. M. Philbin, E. Meier, S. Huffman and P. Boverie, A survey of gender and learning styles, *Sex Roles*, **32**(7-8), pp. 485–494, 1995.
37. A. Honigsfeld and R. Dunn, High school male and female learning-style similarities and differences in diverse nations, *The Journal of Educational Research*, **96**(4), pp. 195–206, 2003.
38. H. Stadler, R. Duit and G. Benke, Do boys and girls understand physics differently?, *Physics education*, **35**(6), p. 417, 2000.
39. M. Gunawan, N. Suranti, N. Nisrina, R. Ekasari and L. Herayanti, Investigating Students Creativity Based on Gender by Applying Virtual Laboratory to Physics Instruction, *International Conference on Teacher Training and Education 2017 (ICTTE 2017)*, Surakarta, Indonesia, October 7, 2017.
40. A. E. Preston, Leaving science: Occupational Exit from Scientific Careers, Russell Sage Foundation, New York, 2004.
41. N. J. Formosa, B. W. Morrison, G. Hill and D. Stone, Testing the efficacy of a virtual reality-based simulation in enhancing users' knowledge, attitudes, and empathy relating to psychosis, *Australian Journal of Psychology*, **70**(1), pp. 57–65, 2018.
42. R. S. Kennedy, N. E. Lane, K. S. Berbaum and M. G, Lilienthal, Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness, *The International Journal of Aviation Psychology*, **3**(3), pp. 203–220, 1993.
43. J. Brooke, SUS-A quick and dirty usability scale, *Usability Evaluation in Industry*, **189**(194), pp. 4–7, 1996.
44. B. G. Witmer and M. J. Singer, Measuring presence in virtual environments: a presence questionnaire, *Presence Teleoperators Virtual Environment*, **7**(3), pp. 225–240, 1998.
45. S. G. Hart, NASA Task load Index (TLX). Volume 1.0; Paper and pencil package, NASA, Washington, DC, 1986.
46. P. R. Pintrich, D. A. F. Smith, T. Garcia and W. J McKeachie, A manual for the use of the motivated strategies for learning questionnaire (MSLQ), Ann Arbor, The University of Michigan, Michigan, 1991.
47. M. Hamilton, R. Jaradat, P. Jones, E. Wall, V. Dayarathna, D. Ray and G. Hsu, Immersive Virtual Training Environment for Teaching Single and Multi-Queuing Theory: Industrial Engineering Queuing Theory Concepts, *ASEE Annual Conference and Exposition*, Tamapa, FL, June 2018, pp. 24–27, 2018.
48. A. Koch, I. Cascorbi, M. Westhofen, M. Dafotakis, S. Klapa and J. P. Kuhtz-Buschbeck, The neurophysiology and treatment of motion sickness, *Deutsches Ärzteblatt International*, **115**(41), p. 687, 2018.
49. N. Bevan, J. Carter and S. Harker, ISO 9241-11 revised: What have we learnt about usability since 1998?, International Conference on Human-Computer Interaction, Los Angeles, CA, August 2–7, pp. 143–151, 2015.

Vidanelage L. Dayarathna is a PhD student in the Industrial and Systems Engineering Department at Mississippi State University. Vidanelage received a Bachelor's degree in Physical sciences and Post Graduate Diploma in operational Research from Sri Lankan Universities. His main research interests include data analytics, modeling and simulation, system of systems, systems thinking skills, and virtual reality. His publications appeared in different journals such as the *Journal of Mechanical Design*, *Expert Systems with Applications, and International Journal of Procurement Management*. With his fellow researchers Vidanelage received three outstanding awards from the ASEE community in 2018. He is an active member of the Institute of Industrial & Systems Engineering (IISE) and the American Society of Engineering Education (ASEE).

Sofia Karam is a PhD student in Industrial and Systems Engineering at Mississippi State University with a concentration in system of systems and virtual reality educational models. Her dissertation topic about the development of an immersive virtual-reality module to assist with big data analytics and decision-making. Before joining our team, Sofia obtained a bachelor's degree in Aerospace Engineering at the International University of Rabat (UIR), Morocco. Also, she earned a Master of Science in Aerospace Engineering with a concentration in fluid flows and jet noise from Mississippi State University. Sofia is an active member of Institute of Industrial and Systems Engineering (IISE).

Raed Jaradat, PhD is an Associate Professor of Industrial and Systems Engineering Department at Mississippi State University and a visiting professor working with the Institute for Systems Engineering Research/MSU/U.S. Army Corps of Engineers. Dr. Jaradat received a PhD in Engineering Management and Systems Engineering from Old Dominion University in 2014. His main research interests include systems engineering and management systems, systems thinking and complex system exploration, system of systems, virtual reality and complex systems, systems simulation, risk, reliability and vulnerability in critical infrastructures with applications to diverse fields ranging from the military to industry. His publications appeared in several ranking journals including the IEEE Systems Journal, and the Computers & Industrial Engineering Journal. His total awarded projects exceed $ 6 M including National Science Foundation (NSF), Department of Defense (DOD), Industry, and other Research Laboratories. Dr. Jaradat's work has been recognized in the IISE professional communities. He received three international awards from the 2018 ASEE National Conference including Industrial Engineering Division's (IED) Best Paper Award, the New IE Educator Outstanding Paper Award, and best paper for the Professional Interest Council 1 (PIC 1). Dr Jaradat is elected as a Director of Networking and Outreach for the Modeling and Simulation (M&S) Board of Directors at The Institute of Industrial and Systems Engineers (IISE). Dr. Jaradat also serve as a Guest Lead Editor for the IEEE Transactions on Engineering.

Michael A. Hamilton, PhD is an Associate Director at Mississippi State Institute for System Engineering Research (ISER) in Vicksburg, MS. He received his Doctorate, Master and Bachelor degrees in Industrial and Systems Engineering from Mississippi State University and has a graduate certificate in Modeling, Simulation, and Visualization Engineering from Old Dominion University. Currently, he is the technical lead for Big Data Analytic and Visualization, and Surrogate Modeling efforts in conjunction with the U.S. Army Engineer Research and Development Center (ERDC). Over the past 4 years, Dr. Hamilton have been leading the efforts for developing immersive virtual environments for conducting data analyzes of tradespace data sets. The immersive data visualization systems allows stakeholders the ability to visualize the tradespace options, subset the data, and work in collaboration with other analysts within the same virtual environment. Other research interest includes using virtual reality for enhancing classroom education in engineering programs. Dr. Hamilton and his colleagues was awarded Industrial Engineering Division Best Paper, Industrial Engineering Education Outstanding Paper, and Professional Interest Council 1 Best Paper at the 2018 American Society for Engineering Education Conference for their research in using virtual reality to assisting students with understanding the concepts of queuing theory.

Morteza Nagahi is a doctoral candidate and graduate research assistant at the Department of Industrial and Systems Engineering, Mississippi State University. Previously, Morteza received a bachelor's degree in Mechanical Engineering from the University of Tehran, and a joint MBA degree specialized in Finance and Marketing from Mazandaran UST and IIT Madras in 2012 and 2014, respectively. Morteza's works have been published in prestigious journals including, *Expert Systems With Applications*, *Engineering Management Journal*, *International Journal of System of Systems Engineering*, *Journal of Computational Design & Engineering*, and *International Journal of Procurement Management*. Moreover, he is a reviewer in several journals and conferences, including IEEE TEM, IEEE Systems, Systems Engineering, IEEE VR, ASEE, etc. He is a member of ASEE, IEEE, ASEM, INFORMS, INCOSE, and IISE. His main areas of research interest are systems thinking, complex systems, engineering education, individual differences, and advanced statistical analysis.

Sayali Joshi is a PhD student and a research assistant at the Department of Industrial and Systems Engineering at Mississippi State University. She has pursued a bachelor's in Statistics followed by a master's in actuarial science from Mumbai, India. From 2017, she is working on a dual degree with masters in statistics and from 2018 she applied for a dual degree program in PhD in Industrial and Systems Engineering. Sayali was awarded, PCI Daniel P. Jenny fellowship for 2018–2019, where she developed a Virtual Reality module and conducted research for safety improvement of precast/prestressed industry employees. Her major research areas include fuzzy analysis, Bayesian network, and various statistical analysis.

Junfeng Ma, PhD, his expertise locates in sustainability driven logistics network modeling and optimization, process monitoring and anomaly detection. He has published more than 40 peer-reviewed publications and led several collaborative research projects. Recent research efforts include developing olefin supply chain network given seasonality supplies and stochastic consumer demands, integrating sustainability consideration into vehicle/vessel routing to collect and remove garbage, and recognizing appropriate locations and allocating resources to satisfy customers' stochastic demands. He got dual title PhD degree in Industrial Engineering and Operations Research from The Pennsylvania State University in 2016. He is an active member in INFORMS, ASME and IISE.

Bouteina Driouche is a PhD candidate in the Department of Aerospace Engineering at Mississippi State University. Her dissertation revolves around developing Markov chain models that predicts performance degradation issues of military vehicles in order to provide insight for the vehicle maintainer of vehicle performance issues before a critical component failure occurs. Bouteina obtained a "Diplôme d'Ingénieur d'Etat" in Aerospace and Aeronautical Engineering at the International University of Rabat (UIR), Morocco. In addition, she earned a Master of Science in Aerospace Engineering at Mississippi State University. She performed an investigation of the effectiveness of polymer-based materials for radiation shielding of flight vehicles for her master's thesis. Bouteina currently works as a Graduate Research Assistant at the Institute for Systems Engineering Research (ISER).

Omar Ashour, PhD is an Associate Professor of Industrial Engineering at Pennsylvania State University, The Behrend College. Dr. Ashour received the BS degree in Industrial Engineering/Manufacturing Engineering and the MS degree in Industrial Engineering from Jordan University of Science and Technology (JUST) in 2005 and 2007, respectively. He received his M.Eng. degree in Industrial Engineering/Human Factors and Ergonomics and the PhD degree in Industrial Engineering and Operations Research from Pennsylvania State University (PSU) in 2010 and 2012, respectively. Dr. Ashour was the inaugural recipient of William and Wendy Korb Early Career Professorship in Industrial Engineering in 2016. Dr. Ashour's research areas include applied decision making, modeling and simulation, virtual reality, and process improvement. He contributed to research directed to improve engineering education.