Sensor-based Modeling of Problem-Solving in Virtual Reality Manufacturing Systems

Rui Zhu
School of Industrial and Systems Engineering
The University of Oklahoma, Norman, OK, 73019, USA
rui.zhu@ou.edu

Faisal Aqlan*
Department of Industrial Engineering
University of Louisville, Louisville, KY, 40292, USA
faisal.aqlan@louisville.edu

Richard Zhao
Department of Computer Science
University of Calgary, Calgary, Alberta, T2N 1N4, Canada
richard.zhao1@ucalgary.ca

Hui Yang
Complex Systems Monitoring, Modeling and Control Laboratory
The Pennsylvania State University, University Park, PA, 16802, USA
huy25@psu.edu

ABSTRACT

Problem-solving is the process to achieve a goal when the solution path is uncertain. Recently, technological advancements have changed problems' characteristics and their solutions in engineering fields. Strong problem-solving skills are essential to allow engineers to assess new problems and quickly implement solutions. Engineering problem-solving skills are first educated in schools and usually evaluated through written exams. However, high grades in exams do not represent sufficient problem-solving skills in real-world engineering problems. Decision making with insufficient problem-solving skills in real world may result in costly consequences. Therefore, it is imperative to evaluate and reinforce problem-solving skills of engineering students in real-world problems. With the rapid technological advancements, availability of virtual reality (VR) and eye-tracking facilitates the study of engineering problem-solving. The immersive environment created by VR enables students to better understand and solve real-world engineering problems. On the other hand, eye-tracking allows for studying fundamental cognitive processes during information processing. It is critical to integrate VR simulation with data-driven modeling of eye movements to evaluate and enhance engineering problem-solving skills. In this paper, we integrate sensing technology (i.e., eye-tracking) and virtual reality (VR) to model problem-solving in manufacturing systems. A novel data-driven model that integrates signal detection theory (SDT) with Conflict & Error (C&E) is developed to quantify engineering problem-solving skills. First, we simulate a manufacturing system in a VR game environment. Students are given an assembly problem to produce a car toy that

* Corresponding author

satisfies some particular requirements in the VR manufacturing system while eye-tracking data are collected throughout the assembly process. Second, eye-tracking data are analyzed with a SDT model to quantify problem-solving skills. Third, a joint SDT-C&E model is developed to analyze eye-tracking data and benchmark with results generated from the SDT model. Experimental results show that the joint SDT-C&E model is more effective to quantify problem-solving skills of engineering students than the SDT model.

Keywords: *sensor-based modeling, virtual reality, eye-tracking, signal detection theory, conflict and error, problem-solving, manufacturing*

1. INTRODUCTION

Problems are defined as discrepancies between initial problem states and goal states (Ward, 2012). Problem-solving is a cognitive process that finds solution paths to achieve the goal state (Wang and Chiew, 2010). Recent technological advancements have changed the characteristics of problems and their solutions, especially in engineering fields (Autor and Price, 2013). Today, engineers need to solve problems that they have never experienced before. Strong problem-solving skills are essential to allow engineers to assess new problems and quickly implement solutions. Engineering problem-solving skills are first educated in schools and evaluated through written exams. However, high grades in exams do not represent sufficient problem-solving skills in real-world engineering problems (Jonassen et al., 2006). Furthermore, decision making with insufficient problem-solving skills in real world may result in costly consequences. Therefore, it is imperative to evaluate and reinforce problem-solving skills of engineering students in real-world problems.

Universities may not be able to introduce the latest manufacturing systems and technologies into the learning factory due to limited resources. In addition, manufacturing safety is important to reduce the risks of workplace injury. Injuries to students may cause significant compensation and medical treatment costs (Wang et al., 2018). Therefore, it is imperative to develop a cost-effective and safe learning environment for engineering students to get hands-on training. With the rapid technological advancements, availability of virtual reality (VR) and eye-tracking facilitates the study of engineering problem-solving. VR simulates a real-world experience in an immersive virtual environment, which allows users to interact with virtual objects and immerse in the 3D simulation. This enables students to better understand real-world engineering problems and solve them. On the other hand, eye-tracking is an emerging technology that allows eye movements to be monitored. An operational definition of eye-tracking by Poole and Ball (2006) states that eye-tracking is a technique whereby an individual's eye movements are measured so that the researcher knows both where a person is looking at any given time and the sequence in which their fixations are shifting from one location to another. Studies show that increased eye movements reveal the increment of cognitive activities. For example, the large number of fixations and saccades reveal human subjects' attention on relevant stimuli (Zagermann, Pfeil, and Reiterer, 2018). Thus, eye-tracking allows for revealing cognitive activities and studying the fundamental cognitive processes during information processing. Problem-solving, as a fundamental cognitive process (Wang and Chiew, 2010), which can be understood by utilizing eye-tracking technology and thereby evaluate problem-solving skills. Therefore, it is critical to integrate VR simulation with data-driven modeling of eye movements to evaluate and enhance engineering problem-solving skills.

In this paper, we develop an analytical model that integrates signal detection theory (SDT) with Conflict & Error (C&E) to quantify problem-solving skills of engineering students. First, we simulate a manufacturing system in a VR game environment. Students are given an assembly problem to produce a car toy that satisfies some particular requirements in the VR manufacturing system while eye-tracking data are collected throughout the assembly process. Second, eye-tracking data are analyzed with a SDT model to quantify problem-solving skills. Third, we develop a joint SDT-C&E model to analyze eye-tracking data and benchmark with results generated from the SDT model.

Experimental results show that the proposed joint SDT-C&E model is more effective to quantify problem-solving skills of engineering students than the SDT model.

2. RELEVANT LITERATURE

The ongoing globalization and advancements in technology confront people with complex environments that demand numerous problems to be solved (Fischer et al., 2012). The ability to solve such problems is an essential competence and is required for active participation in today's society (Eichmann et al., 2019). Problem-solving is the basis of many scholastic learning processes and is therefore regarded as a fundamental goal of education (OECD, 2013). A study concludes that problem-solving skills are more important than numerical or communication skills for a worker to be successful in the workplace (Felstead et al., 2013). Because of the changes in the characteristics of engineering problems and solutions, most engineering problems are open-ended (Belski, 2011; Mourtos et al., 2004). These problems often possess vaguely defined goals, multiple solutions, and multiple criteria for evaluating the solutions. Due to the complexity of real-world engineering problems, learning from schools does not adequately prepare engineering students to solve real-world problems. Therefore, it is significant to bridge the gap between textbook theory and real-world application for engineering students.

This research utilizes SDT (Green and Swets, 1966) from the psychology literature, where the presence or absence of events are used to analyze behaviors. We utilize SDT to represent whether a student's choice matches that of a subject matter expert (SME). While SDT provides us with the analysis of errors, we also integrate the concept of conflict from neuropsychology. We use conflict to represent the discrepancies between a student's gaze and the student's subsequent choice. We detail SDT model in section 3.2 and SDT-C&E model in section 3.3.

2.1 VR Simulation of Problem-Solving

VR simulations provide multi-dimensional human experiences which mimic real-world experiences. In the past few years, researchers have given more attention to the application of VR simulation in different areas including human computer interaction (Arora et al., 2019), sports (Macedo et al., 2019), biology (Desmeulles et al., 2006), education (Beck, 2019), smart manufacturing (Yang et al., 2019), and problem-solving (Hwang and Hu, 2013). VR is a useful tool for teaching problem-solving skills, especially when it is difficult to perform the task in real-life. For example, simulating a car factory in VR can help students learn problem-solving skills for manufacturing processes without the need to visit a physical plant (Aqlan et al., 2020). In VR, 3D objects and environments can be created which allows learners to interact and appeal to their visual or other senses. A study by Hwang and Hu (2013) developed a VR learning environment to study the peer learning behaviors and their impacts on geometry problem-solving. The utilization of VR allowed for synchronous manipulation of objects and communication among multiple users. It also improved the problem-solving skills for the participants. However, this study only utilized questionnaire and interviews to collect data about the student behavior in the VR environment. One advantage of using VR is the ability to collect data from participants via sensing technology, which can provide valuable insights about user behavior and problem-solving skills. Another study by Jin and Lee (2019) compared problem-solving styles between desktop and VR environments based on the influence of design tools in ideation.

The study found that VR can provide frequent modifications of solutions and high manipulability of user interface. However, the study argued that higher usability does not always produce desired outcomes. This can be addressed by developing effective VR environments that are equivalent to the actual environment as well as taking into consideration the similarity between the tasks performed in both environments. In Tang et al. (2012), a VR theme-based game was developed to replace traditional laboratory activities in electrical and computer engineering. The game was designed with specific considerations of the nature of problem-solving in the manufacturing context. Students needed to provide solutions with their Hardware Description Language code. They were allowed to debug the code until the problems were solved, for example, fixing a malfunctioning traffic light. In this way, students were able to implement their domain knowledge and improve problem-solving skills. The game aimed at providing a fun learning environment to promote strategic problem-solving. In Man et al. (2013), VR-based training programs have also been used to strengthen problem-solving skills of people with traumatic brain injury, so as to enhance their employment opportunities. The problem-solving skills in their research were measured by Wisconsin Card Sorting Test, Tower of London Test, and Vocational Cognitive Rating Scale. In manufacturing education and industry, VR can be used to provide more efficient ways to solve problems and improve design choices. According to Milella (2015), when compared to traditional desktop-based modelling and simulation tools, VR offers unquestionable advantages in terms of rapid problem-solving. The author suggests that further research is required to develop more efficient VR simulations for manufacturing, as well as to evaluate time and cost saving in comparison with desktop-based modelling and simulation tools. Hence, there is an urgent need to leverage large amounts of sensing data and analytical models to quantify training outcomes in VR simulations. To address these challenges, this research develops a VR simulation for manufacturing environments to evaluate and quantify problem-solving skills. The research utilizes sensing technology for data collection and physiological theories to analyze the collected data.

2.2 Eye-Tracking for Problem-Solving

Problem-solving requires many cognitive processes. Eye-tracking has been shown to effectively reveal cognitive activities (Eckstein et al., 2017) and has been utilized to examine students' visual attention while solving multi-choice science problems (Tsai et al., 2012). It is found that students pay more attention to the choice they prefer rather than alternatives they reject and spend more time on inspecting relevant factors than irrelevant ones in problems. A study utilized eye-tracking to examine how the problem-solving performance of learners varies with different levels of prior knowledge (Lee et al., 2019). The authors derived multiple performance aspects, such as accuracy in visual attention and cognitive load, which are possibly affected by prior knowledge. The study employed a medical simulation game to empirically examine whether the level of prior knowledge affects those performance aspects. Research on eye-tracking is increasing owing to its ability to facilitate many different tasks (Klaib et al., 2021). Availability of eye-tracking data has been shown to facilitate the study of problem-solving. However, very little has been done to quantify problem-solving skills with eye-tracking data analytics. In this paper, data-driven models of eye movements are developed to quantify the problem-solving skills in the VR environment.

3. RESEARCH METHODOLOGY

The proposed research methodology integrates VR simulation and eye-tracking to study manufacturing problem-solving and develop analytical models for measuring students' performance. As shown in Figure 1, the first step is to develop physical simulations about the assembly of physical car toys, which are integrated into an undergraduate course on "manufacturing systems". The physical simulations form the basis for developing the VR simulation. Eye-tracking is integrated with VR simulation to collect data on the problem-solver's performance, and the data is analyzed using SDT and C&E models. Heat maps are then developed to visualize the performance of the problem-solvers in terms of selection of the car toy components. Other data are collected from the VR simulation including weight and price of the product and user switches between the assembly stations. The data is visualized on a radar chart based on a composite index that represents the overall performance of the problem-solver.

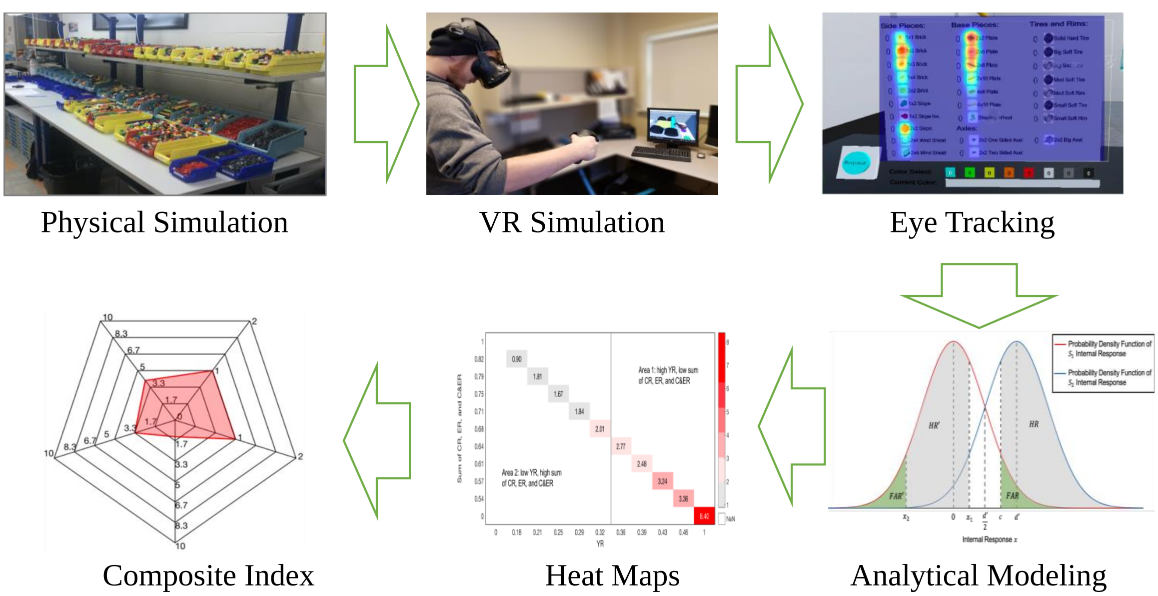

Figure 1. Research methodology

3.1 Virtual Reality Simulation of Manufacturing System

In this paper, we simulate a manufacturing system in VR to evaluate and enhance the problem-solving skills of engineering students. The VR manufacturing system was developed in the Unity game engine and equipped with an HTC Vive VR headset, wireless controllers, and base stations for motion tracking (Zhao et al., 2019). Students wore a VR headset and saw through the headset a virtual manufacturing system composed of a series of workstations. The students were able to interact with objects in the virtual environment using the wireless controller. The VR headset was integrated with Tobii eye-tracking technology, allowing the system to identify coordinates and objects that students were looking at, at any given time during the simulation.

In the VR manufacturing system, students were asked to assemble a car toy that satisfied some particular customer requirements. Students were first presented with audio instructions on how to interact with the virtual manufacturing system. Once students felt comfortable, they could press a button to start the manufacturing assembly process. There were seven stations in the virtual manufacturing system. Students were allowed to switch between stations at any time. The first station was a *requirement* station, where students were given a set of customer requirements as shown in Table 1. After students read the requirements, they moved to the next station, the *component selection* station. This station includes the selection board, which is the area of interest (AOI) for eye-tracking in the VR simulation. The component selection station is shown in Figure 2 (a). Components were selected when a student pointed at them and pressed a trigger on the wireless controller. Each component came with a selection of 8 colors. After students selected the components they desired, they moved over, in order, to the *base* station, *wheel and axle* station, *tire and* *rim* station, *sides* station, and *roof* station as shown in Figure 2. Students went through each station to assemble car toys. Unused components were put into a red trash box. Each component was associated with a weight and a cost, which simulated the real-world manufacturing process where materials had associated weights and costs. Total weight and cost of the car toy were displayed in the VR environment. Once students completed the assembly of car toy, they pressed a "finish" button and the simulation stopped.

Eye-tracking and trace data, including students' choices of components, switches between stations, durations, frequencies, and coordinates of fixations, were collected along the assembly process. Eye-tracking and trace data of students were compared with the data of a SME, a person whose knowledge was accepted as the gold standard and sets the expert criterion for the assembly process of the car toy, to quantify the problem-solving skills of students. In our study, SME is the project lead who has extensive experience conducting problem-solving training for students and professionals.

Table 1. Customer requirements for the car toy

Vehicle Requirements	Functional Requirements
(a) Vehicle weight is between 20 and 30 grams (b) Material cost <= $9 (c) Vehicle must have four tires (with axles), wind shield, driver, steering wheel, and roof (d) All tires must be small soft (e) Vehicle base width and length are 4 dots and 6 dots, respectively (f) Vehicle must fit completely within the design footprint "parking space" (g) Number of different colors for plastic blocks >= 5 (excluding driver and wind shield)	(a) Driver must be able to get in and out of the vehicle and see where he is going while traveling (b) Vehicle must be able to travel over ramp conditions, stay on ramp, and cross the finish line fully intact (c) Vehicle must remain intact following a drop test

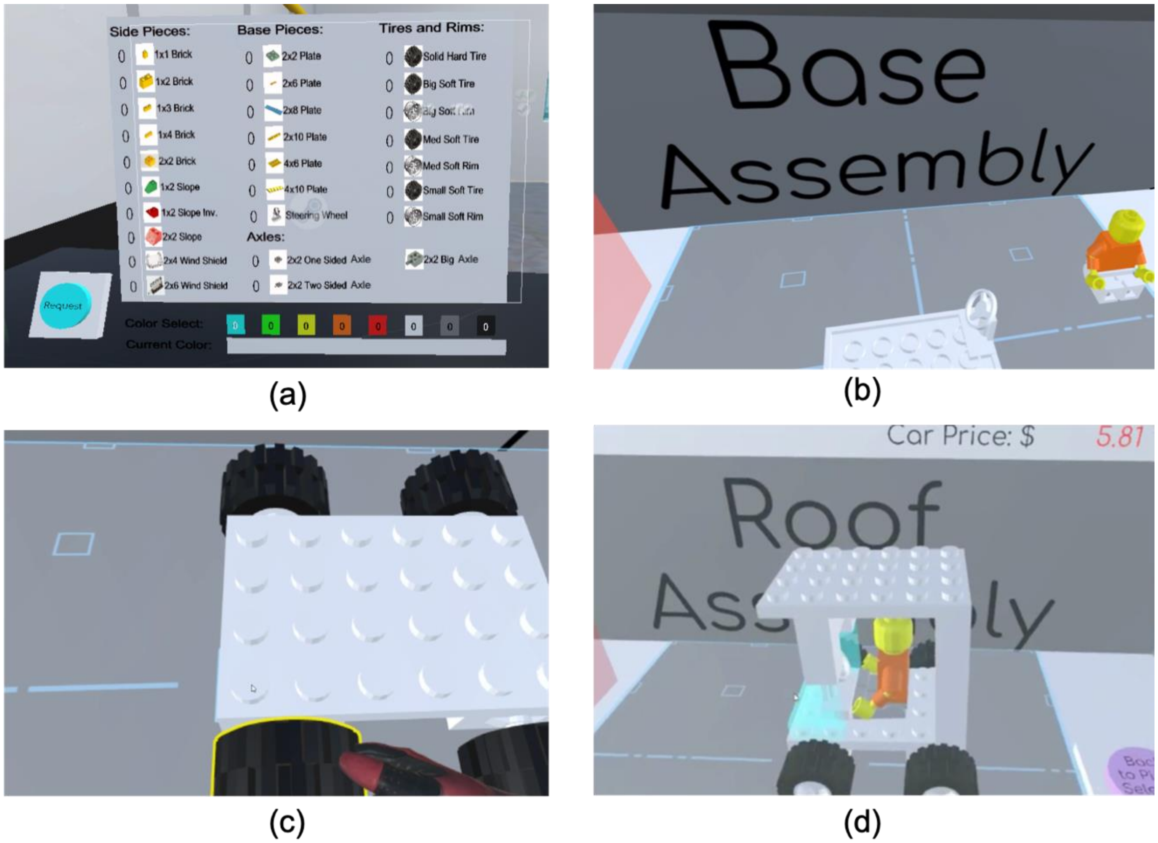

Figure 2. Example stations in the virtual manufacturing system: (a) component selection station; (b) base station; (c) tire and rim station; (d) roof station.

3.2 Signal Detection Theory (SDT)

SDT has been widely applied in areas where two different stimuli must be discriminated. In SDT, every time stimulus S_1 or S_2 is shown, a student generates an internal response x in his/her mind. This internal response x is drawn from a normal distribution with mean μ and standard deviation σ. When the stimulus is absent (i.e., stimulus S_1), $\mu = 0$ and $\sigma = 1$, that is, S_1 internal response follows a standard normal distribution $N(0,1)$. The cumulative distribution function of S_1 internal response is denoted as Φ_0. When the stimulus is present (i.e., stimulus S_2), S_2 internal response is normally distributed with a mean d' and a standard deviation σ. The cumulative distribution function of S_2 internal response is denoted as $\Phi_{d',\sigma}$. For the sake of simplicity,

S_1 and S_2 internal responses are often assumed to have the same standard deviation $\sigma = 1$ (Barrett et al., 2013).

The ability of a student to discriminate stimulus S_1 from S_2 depends on the extent that S_1 and S_2 internal responses in the student's mind are distinguishable. A larger separation between S_1 and S_2 internal responses represents a better sensitivity for discriminating stimuli S_1 and S_2. Therefore, d' is also called a sensitivity index.

<u>*Definition 1*</u> Two possible stimuli S_1 and S_2 are defined as "signal absent" and "signal present", respectively. Four possible outcomes are defined depending on the internal responses of students to stimuli (see Table 2).

Table 2. SDT possible outcomes

Response / Stimulus	S_1 internal response	S_2 internal response
S_1	Correct rejection	False alarm
S_2	Miss	Hit

Hit rate (HR) is the probability of responding S_2 internally in student's mind when the signal is present. As shown in Figure 3, HR is calculated as the area under probability density function of S_2 internal response that exceeds a decision criterion c. Students respond S_1 if internal response $x \leq c$ and respond S_2 if $x > c$. Cumulative distribution function of a normal distribution with mean μ and standard deviation σ evaluated at x is:

$$\Phi(x, \mu, \sigma) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{1}$$

HR is derived from equation (1) as:

$$\mathrm{HR} = 1 - \Phi(c, d', \sigma = 1) \tag{2}$$

Similarly, false alarm rate (FAR) is formulated as:

$$\text{FAR} = 1 - \Phi(c, 0, \sigma = 1) \quad (3)$$

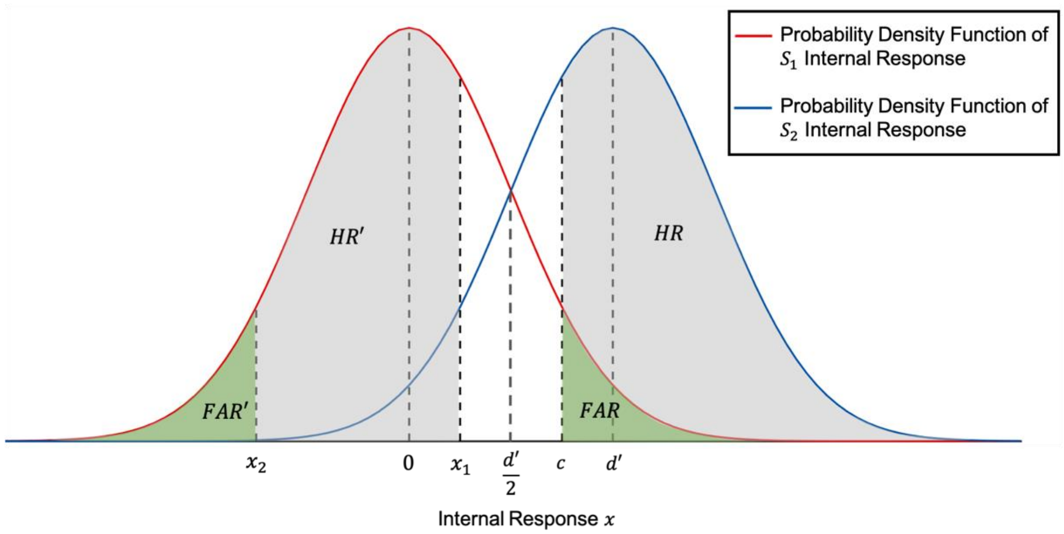

Figure 3. Derivation of sensitivity index d'. HR and HR' are symmetric with respect to $x = \frac{d'}{2}$; FAR and FAR' are symmetric with respect to $x = 0$.

According to Figure 2,

$$x_1 = \Phi_0^{-1}(HR) \quad (4)$$

$$x_2 = \Phi_0^{-1}(FAR) \quad (5)$$

$$x_1 = d' - c \quad (6)$$

$$x_2 = -c \quad (7)$$

$$d' = x_1 - x_2 = \Phi_0^{-1}(HR) - \Phi_0^{-1}(FAR) \quad (8)$$

$$c = -x_2 = -\Phi_0^{-1}(FAR) \quad (9)$$

In this paper, SME's choices of components to assemble the car toy are taken as stimulus S_2. Choices of each student are internal responses. As shown in Table 3, if a component is chosen, then a "Yes" is given to its corresponding "Choice" column, otherwise, a "No" is given. Hits and false alarms are defined based on students' internal responses to S_2 stimulus, i.e., SME's choices.

Table 3. Examples of a hit and a false alarm

Component	Duration (hh:mm:ss)	Frequency	Student Choice	SME Choice	Outcome
1x2Brick	00:00:39	67	Yes	Yes	Hit
2x2Brick	00:00:05	24	Yes	No	False alarm

Definition 2 Hit is defined as each time student matches the choice of SME, i.e., student chooses a component that SME also chooses. False alarm is defined as each time that a student chooses a component that SME does not choose. HR and FAR are formulated as:

$$HR = \frac{\text{No.of hits}}{\text{No.of "Yes" in SME's choices}} \tag{10}$$

$$FAR = \frac{\text{No.of False alarms}}{\text{No.of "No" in SME's choices}} \tag{11}$$

Definition 3 A process, P_i, consists of a set of tasks, T_{ij}, which are performed by students. In this paper, P_1 = design, P_2 = sourcing, P_3 = assembly, and P_4 = inspection.

Therefore, measure of problem-solving skills d' is formulated as:

$$d' = \sum_{i=1}^{P}\sum_{j=1}^{T}\left[\Phi_0^{-1}\left(HR_{ij}\right) - \Phi_0^{-1}(FAR_{ij})\right] \tag{12}$$

Perfect rates which result in infinite $\Phi_0^{-1}\left(HR_{ij}\right)$ and $\Phi_0^{-1}(FAR_{ij})$ are remedied with $1/2N$ rules, specifically, rates of 0 are replaced with $0.5/N$, and rates of 1 are replaced with $(N - 0.5)/N$, where N is the number of “Yes” or “No” in SME’s choices (Stanislaw and Todorov, 1999).

3.3 Joint SDT-C&E Model

In neuropsychology, conflict and error are two important concepts that should be considered in complex problem-solving. Error detects deviation between intentions and actions. Conflict is defined as competition between two or more simultaneously activated response tendencies. However, SDT focuses on measuring error because it analyzes the actions taken by students and fails to measure the conflict of response tendencies in problem-solving. In this paper, we propose to quantify the problem-solving skills of engineering students with a joint SDT-C&E model.

Assumption Choices and decisions of SME are used as the benchmark of student's problem-solving performance.

Definition 4 Four possible outcomes are defined based on students' fixations and choices of car toy components. Correct Choice is defined as cases when student matches the choice of SME and looks at a component if the choice is "Yes" and does not look at a component if the choice is "No". Conflict is defined as cases when student matches the choice of SME, but does not look at a component if the choice is "Yes" and looks at a component if the choice is "No". Error is defined as cases when student does not match the choice of SME and looks at a component if the choice is "Yes" and does not look at a component if the choice is "No". C&E is defined as cases when student does not match the choice of SME and does not look at a component if the choice is "Yes" and looks at a component if the choice is "No". Rates of four outcomes are formulated as:

$$Correct\ Choice\ Rate\ (YR) = \frac{\text{No.of correct choices}}{\text{No.of SME's choices}} \tag{13}$$

$$Conflict\ Rate\ (CR) = \frac{\text{No.of conflicts}}{\text{No.of SME's choices}} \tag{14}$$

$$Error\ Rate\ (ER) = \frac{\text{No.of errors}}{\text{No.of SME's choices}} \tag{15}$$

$$Conflict\ \&\ Error\ Rate\ (C\&ER) = \frac{\text{No.of C\&Es}}{\text{No.of SME's choices}} \tag{16}$$

The proposed SDT-C&E model is formulated as:

$$\rho' = \sum_{i=1}^{P} \sum_{j=1}^{T} \left[\Phi_0^{-1}(YR_{ij}) - \left(\Phi_0^{-1}(CR_{ij}) + \Phi_0^{-1}(ER_{ij}) + \Phi_0^{-1}(C\&ER_{ij}) \right) \right] \tag{17}$$

ρ' is a measure of problem-solving skills. In order to maximize ρ', students need to maximize YR and minimize other rates. Perfect rates are also remedied with $1/2N$ rules.

An Illustrative Example: In order to develop the C&E model, we will first compare student responses to that of SME. This person has an expert level of understanding of the process and knows the best way to solve the problem. His eye tracking fixations and durations will determine the areas of interest to which the students' performance can be compared. Suppose the student needs to select one of two components, Option A or Option B, for the car toy (see Figure 4). Assuming the correct option is A, the number of times a student matches SME performance, in terms of eye tracking and final selection will comprise a correct choice (Y). In Table 4, the SME looked at option A for 30 seconds with two fixations and then selected option A. This is defined as *Y*. Student 1 looked at option A for 16 seconds then option B for 10 seconds then option A again for 15 seconds. Hence, the student looked at option A for 31 seconds with two fixations and then selected option A. This is also defined as *Y*. Student 2, however, looked at option B for 29 seconds and with two fixations and then selected option B. This is defined as an error (*E*). Student 3 looked at option B but selected option A. This is defined as a conflict (*C*). Finally, student 4 looked at option A but selected option B. This is both a conflict and an error (*C&E*).

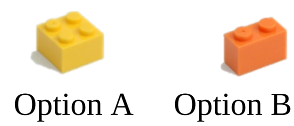

Figure 4. A simple selection problem

Table 4. Comparing students' responses to SME's response

Team Member	Duration (sec.)	No. of Fixations	Final Choice	C&E
SME	**30 → A**	**2**	**A**	**Y**
Student 1	16 → A 10 → B 15 → A	1 1 1	A	Y
Student 2	29 → B	2	B	E
Student 3	20 → B 10 → B	1 1	A	C
Student 4	30 → A	2	B	C&E

4. EXPERIMENTAL RESULTS

In the experiment, we collected the data of 24 undergraduate engineering students and 1 SME in the United States through a user study. All 24 students were undergraduate engineering students from a public university in the United States. The average age was 18 years. The students were recruited from several introductory undergraduate engineering classes. Participation in the experiment was completely voluntary, and students could withdraw from the experiment at any time. Participants were provided $50 gift cards for their involvement in the study. The study was approved by the Institutional Review Board (IRB) of the university. Figure 5 shows the experimental setup.

Students and SME completed the VR simulation as described in Section 3.1. The average weight of car toys is 29.36g with a standard deviation of 13.37g. The average time the participants spent on the assembly process is 16.08 minutes with a standard deviation of 4.94 minutes. Students' eye-tracking and trace data were collected as they went through the simulation. In order to measure the performance of assembly tasks in virtual manufacturing systems, we design a VR-based composite index which involves consideration of cycle time, number of station switches, weight, price, and quality of car toys. The following sections describe the VR-based composite index in detail. The proposed SDT-C&E model is evaluated and validated by comparing correlations between VR-based composite index and measures of engineering problem-solving skills.

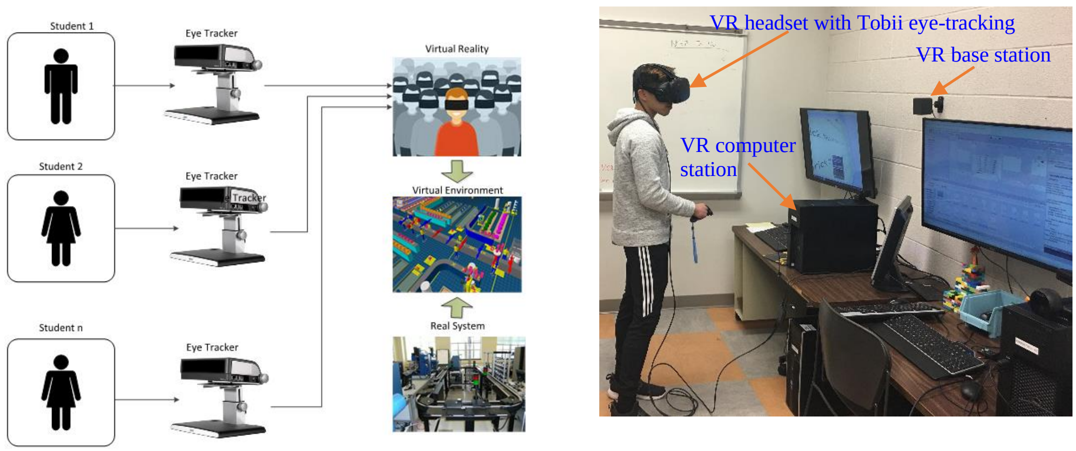

Figure 5. Experimental setup (left) and a student running the VR simulation (right)

4.1 Sensitivity Index d' in SDT Model

Eye-tracking data are analyzed with the SDT model to quantify engineering problem-solving skills. Figure 6 shows the sensitivity index d' of engineering students. Sensitivity index d' increases as HR increases. HR of engineering students ranges from 0.5 to 1. Sensitivity index d' decreases as FAR increases. FAR of engineering students has a range of 0 to 0.72. The zoomed-in figure is equally divided into four areas. Note that engineering students in Area 1 have the highest d' values compared with Areas 2, 3, and 4, because they have high HR and low FAR. Engineering students in Area 4 have the lowest d' values compared with other 3 areas, because they have low HR and high FAR. d' ranges from 0 to 3.83 where SME has the highest d' value of 3.83 with HR of 1 and FAR of 0, which suggests that SME has better problem-solving skills than engineering students. Conflict and error are two important concepts in complex problem-solving. SDT model is effective to quantify the problem-solving skills in terms of the error of responses. However, it fails to account for conflict of response tendencies.

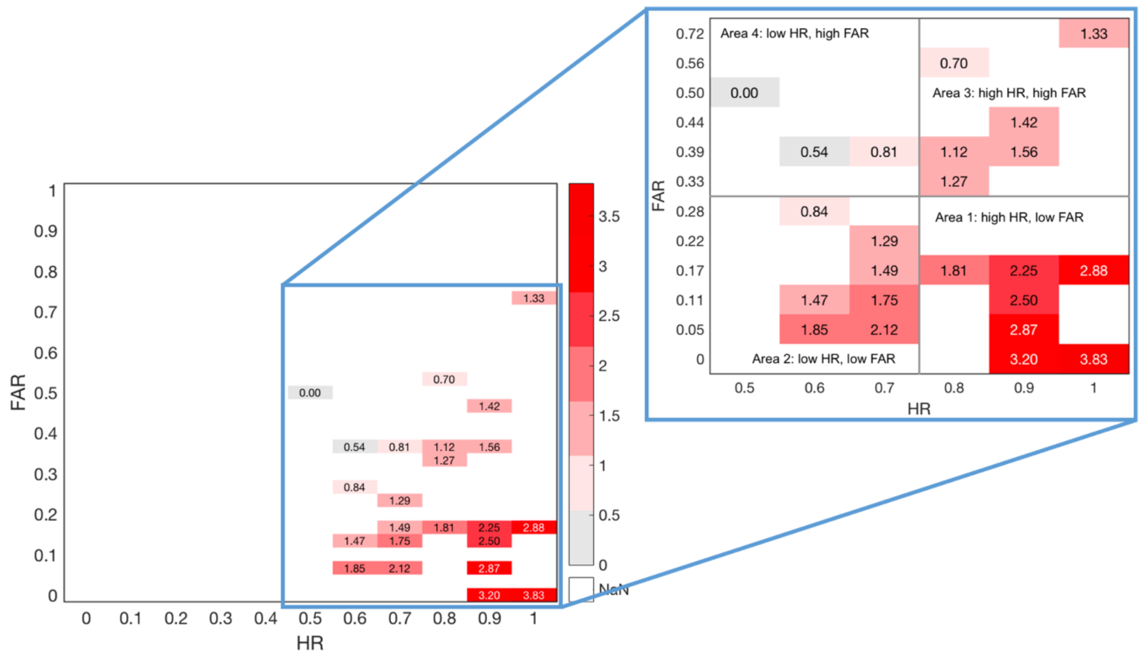

Figure 6. Heatmap of sensitivity index d'.

4.2 Measure of Problem-Solving Skill ρ' in Joint SDT-C&E Model

We use the proposed SDT-C&E model to quantify the problem-solving skills of engineering students. Measure ρ' of engineering students is shown in Figure 7. ρ' increases with the increment of YR. YR ranges from 0.18 to 1. ρ' decreases as CR, ER, or C&ER increases. The sum of CR, ER, and C&ER have a range of 0 to 0.82. Notably, Area 1 has higher ρ' values than Area 2 due to high YR and low sum of other rates. On the other hand, Area 2 has low ρ' values due to low YR and high sum of CR, ER, and C&ER. SME has the highest ρ' value of 8.40, suggesting that SME has the highest level of problem-solving skills. Compared to SDT model, joint SDT-C&E model considers both error of responses and conflict of response tendencies.

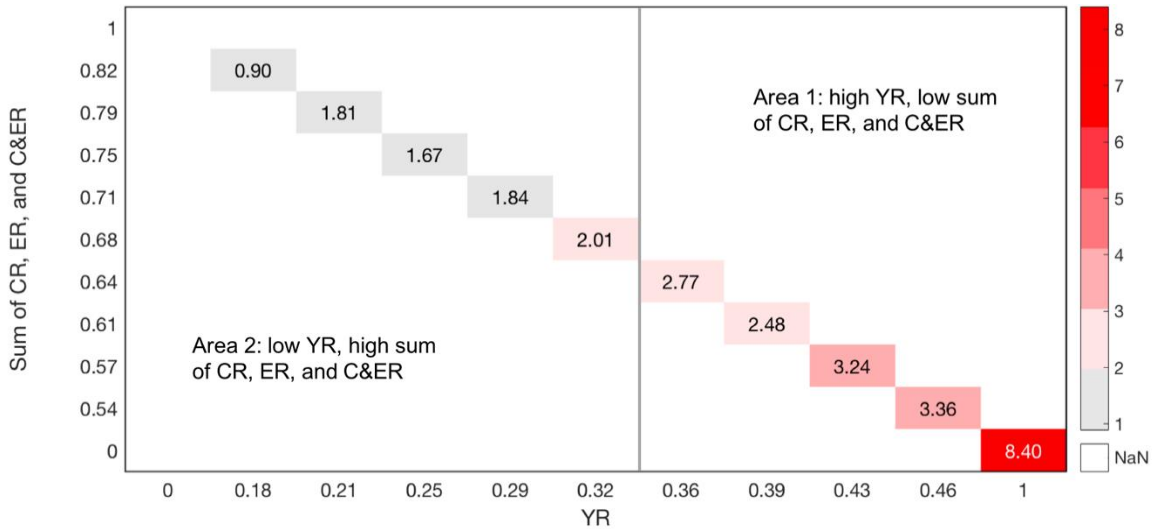

Figure 7. Heatmap of problem-solving skills measure ρ'.

4.3 Correlations between VR-Based Composite Index and Measures of Problem-Solving Skills

Five performance metrics involving cycle time, number of station switches, quality, weight, and price of car toy, are considered in VR-based composite index to measure the task performance of each engineering student. A score is given to each performance metric. Scores of cycle time, number of station switches, and quality of car toy have a range of 0 to 10.

Scores of cycle time and number of station switches are formulated based on the reverse scaling, so that students get low scores if they have long cycle times or large numbers of station switches. Score of cycle time is formulated as:

$$Time\ Score_k = \frac{max(\boldsymbol{CT}) - CT_k}{max(\boldsymbol{CT}) - min\ (\boldsymbol{CT})} \times 10 \tag{18}$$

where k denotes the index of participants, $k = 1, 2, \dots, K$. The total number of participants K is 25 which involves 24 students and 1 SME. $Time\ Score_k$ denotes k^{th} participant's score of cycle time. $\boldsymbol{CT}$ is the set of cycle times $(CT_1, CT_2, \dots, CT_K)$. Score of number of stations switches is formulated as:

$$Switch\ Score_k = \frac{max(\boldsymbol{nSwitch}) - nSwitch_k}{max(\boldsymbol{nSwitch}) - min\ (\boldsymbol{nSwitch})} \times 10 \tag{19}$$

where $Switch\ Score_k$ denotes k^{th} participant's score of number of station switches, $\boldsymbol{nSwitch}$ is the set of numbers of station switches $(nSwitch_1, nSwitch_2, \dots, nSwitch_K)$.

The total score of car toy quality is 10 points. Each violation of customer requirements (e.g., installation of small soft tires, axles, wind shield, steering wheel, roof, base with size of 4×6, driver) deducts 1 point from the starting state. Incompleteness of car toy results in deduction of 0 to 3 points from the starting state. If a car toy meets weight (i.e., between 20 and 30 grams) or price (≤ \$9) requirements, it obtains a score of 2 for weight or price of the car toy. Otherwise, it obtains a score of 1.

Figure 8 demonstrates spider charts of the five performance metrics. Red area in the spider chart represents VR-based composite index of each participant. SME has the highest index among all the participants, which is 85.5951, because SME has full scores on all performance metrics as shown in Figure 8 (a). Figure 8 (b) and (c) give examples of high and low VR-based composite indices, respectively.

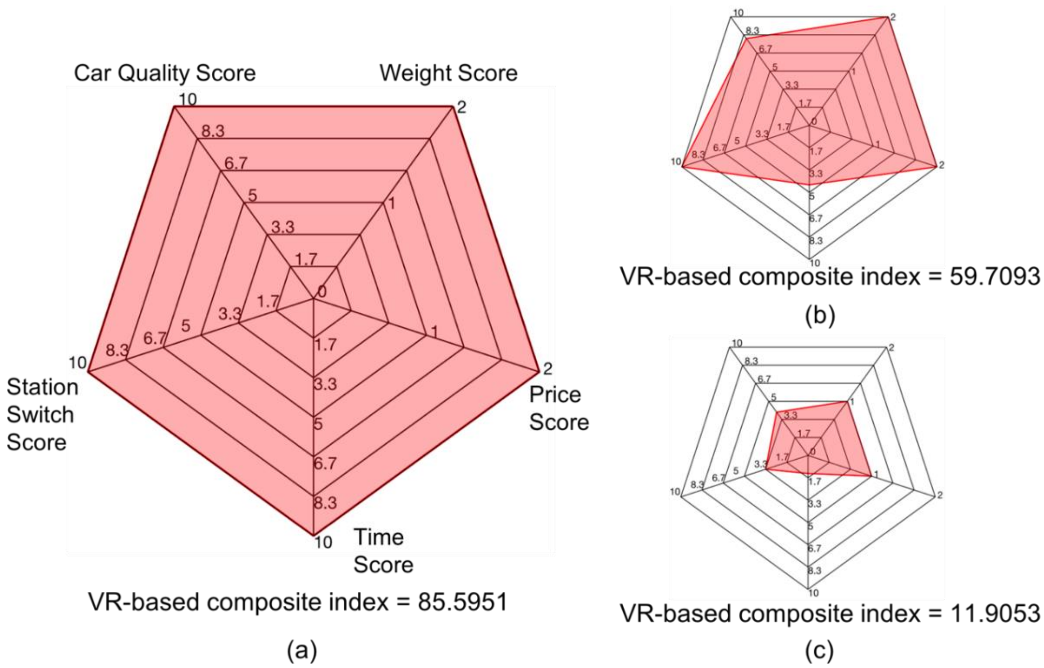

Figure 8. The red spider charts of VR-based composite indices computed from 5 scores, including car quality score, station switch score, scores of car weight, price, and cycle time: (a) Index of SME is 85.5951; (b) An example of high index which is 59.7093; (c) An example of low index which is 11.9053.

We further analyze the correlations between the VR-based composite index and measures of problem-solving skills according to the Pearson's correlation coefficient, which is formulated as:

$$r_{\omega,\tau} = \frac{\sum_{k=1}^{K}(\omega_k - \overline{\omega})(\tau_k - \overline{\tau})}{\sqrt{\sum_{k=1}^{K}(\omega_k - \overline{\omega})^2}\sqrt{\sum_{k=1}^{K}(\tau_k - \overline{\tau})^2}} \tag{20}$$

where ω_k is the measure of problem-solving skill d' or ρ' of k^{th} participant, $\overline{\omega} = \frac{1}{K}\sum_{k=1}^{K}\omega_k$ is the mean of d' or ρ', τ_k is VR-based composite index of k^{th} participant, $\overline{\tau} = \frac{1}{K}\sum_{k=1}^{K}\tau_k$ is the mean of VR-based composite indices.

Table 5. Correlation coefficients between VR-based composite index and measures of problem-solving skills

Correlation Coefficient	τ	τ^2
ρ'	0.6792	0.8219
d'	0.6238	0.6577

Correlation coefficients are summarized in Table 5. ρ' shows higher correlations with both τ and τ^2 than d', which indicates that ρ' is effective in quantifying the problem-solving skills of engineering students by taking conflict and error into consideration. Especially, correlation between τ^2 and ρ' is a lot higher than d', suggesting potential nonlinear correlation between τ and ρ'.

5. CONCLUSIONS

In this paper, we developed an analytical model that integrates SDT with C&E to quantify problem-solving skills of engineering students. First, we simulated a manufacturing system in a VR game environment. Students were given an assembly problem to produce a car toy that satisfied some particular requirements in the VR manufacturing system. Eye-tracking and trace data were collected throughout the assembly process. Second, eye-tracking data were analyzed with a SDT model to quantify problem-solving skills. Third, we developed a joint SDT-C&E model to analyze eye-tracking data and benchmark with results generated from the SDT model.

Experimental results showed that measure of problem-solving skill ρ' generated by the proposed SDT-C&E model had higher correlation with VR-based composite index τ (0.6792 vs 0.6238) and τ^2 (0.8219 vs 0.6577) than sensitivity index d' of SDT model, which suggested that ρ' is effective to quantify engineering problem-solving skills by taking conflicts and errors into account. The higher correlation between ρ' and τ^2 than τ is worth noting because it implies potential nonlinear correlation between measure of problem-solving skills and VR-based composite index. However, a limitation of this work is that the number of participants is small because setting up the equipment and completing the assembly task are time-consuming. In future work, we will recruit more participants in the experiment and collect more data to further validate the model with the iterative design approach and investigate the nonlinear correlation between the measures of problem-solving skills and VR-based composite index.

The VR manufacturing system developed in this paper can serve as a training tool for engineering students to reinforce their problem-solving skills. Additionally, the proposed SDT-C&E model provides a powerful tool to quantify problem-solving skills of engineering students. In this paper, we only compared students' solutions against that of SME. However, given that a problem can have more than one valid solution, this study can be extended by comparing against multiple correct solutions. The SDT-C&E model can be generalized to quantify problem-solving skills in many other disciplines such as healthcare, psychology and cognitive sciences, by comparing one's problem-solving actions with actions of a SME. For example, cardiac surgery requires multiple skills. Novice surgeons can benefit from training on surgical skills utilizing simulation models. However, studies on assessments of training outcome remain sketchy. Current assessments usually rely on subjective observations and logbooks (Lodge and Grantcharov, 2011). If an expert surgeon sets a golden standard with their actions in technical procedures, the proposed model can generate rates of correct choice, conflict, and error for novice surgeons by comparing their actions to the golden standard and then provide assessments of the surgical training outcomes. Further, the rates can help novice surgeons gain insights on how to improve their skills.

ACKNOWLEDGMENT

This research was supported by NSF award # 183074 and supplement # 1905680. Any opinions, findings, or conclusions found in this paper are those of the authors and do not necessarily reflect the views of the sponsor.

NOMENCLATURE

S_1	Stimulus absent
S_2	Stimulus present
x	Internal response to stimulus S_1 or S_2 in student's mind
μ	Mean of internal response distribution. $\mu = 0$ when stimulus is absent; $\mu = d'$ when stimulus is present.
σ	Standard deviation of internal response distribution. for the sake of simplicity, $\sigma = 1$ for both S_1 and S_2 internal response distributions.
Φ_0	Cumulative distribution function of S_1 internal response
$\Phi_{d',\sigma}$	Cumulative distribution function of S_2 internal response
d'	Sensitivity index, serving as a measure of problem-solving skills in SDT model.
c	Decision criterion. Students respond S_1 if internal response $x \leq c$ and respond S_2 if $x > c$.
HR	Hit rate
FAR	False alarm rate
P_i	i^{th} process in the assembly problem
T_{ij}	j^{th} task in Process P_i
YR	Correct choice rate
CR	Conflict rate
ER	Error rate
$C\&ER$	Conflict & Error rate
ρ'	Measure of problem-solving skills in joint SDT-C&E model
k	index of participants, $k = 1, 2, \dots, K$. $K = 25$ involves 24 students and 1 SME.
$Time\ Score_k$	Cycle time score of k^{th} participant
CT	A set of cycle times $(CT_1, CT_2, \dots, CT_K)$
$Switch\ Score_k$	k^{th} participant's score of number of station switches
$nSwitch$	A set of numbers of station switches $(nSwitch_1, nSwitch_2, \dots, nSwitch_K)$
ω_k	Measure of problem-solving skill d' or ρ' of k^{th} participant
$\overline{\omega}$	$\overline{\omega} = \frac{1}{K}\sum_{k=1}^{K}\omega_k$ is the mean of d' or ρ'
τ_k	VR-based composite index of k^{th} participant
$\bar{\tau}$	$\bar{\tau} = \frac{1}{K}\sum_{k=1}^{K}\tau_k$ is the mean of VR-based composite indices
$r_{\omega,\tau}$	Pearson's correlation coefficient between measure of problem-solving skill ω and VR-based composite index τ

REFERENCES

Aqlan, F., Zhao, R., Yang, H. and Ramakrishnan, S., 2020, "A Virtual Learning Factory for Advanced Manufacturing," In Proceedings of the Winter Simulation Conference.

Arora, J., Saini, A., Mehra, N., Jain, V., Shrey, S., and Parnami, A., 2019, "VirtualBricks: Exploring a Scalable, Modular Toolkit for Enabling Physical Manipulation in VR," Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, New York, NY, pp. 1–12.

Autor, D. H., and Price, B., 2013, "The changing task composition of the US labor market: An update of Autor, Levy, and Murnane (2003)," Unpublished Manuscript.

Barrett, A. B., Dienes, Z., and Seth, A. K., 2013, "Measures of metacognition on signal-detection theoretic models.," Psychological Methods, 18(4), pp. 535–552.

Beck, D., 2019, "Special Issue: Augmented and Virtual Reality in Education: Immersive Learning Research," Journal of Educational Computing Research, 57(7), pp. 1619–1625.

Belski, I., 2011, "TRIZ course enhances thinking and problem solving skills of engineering students," Procedia Engineering, 9, pp. 450–460.

Desmeulles, G., Querrec, G., Redou, P., Kerdélo, S., Misery, L., Rodin, V. and Tisseau, J., 2006, "The virtual reality applied to biology understanding: The in virtuo experimentation," Expert systems with applications, 30(1), pp.82-92.

Eckstein, M. K., Guerra-Carrillo, B., Miller Singley, A. T., and Bunge, S. A., 2017, "Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development?," Developmental Cognitive Neuroscience, 25, pp. 69–91.

Eichmann, B., Goldhammer, F., Greiff, S., Pucite, L., and Naumann, J., 2019, "The role of planning in complex problem solving," Computers & Education, 128, pp. 1–12.

Felstead, A., Gallie, D., Green, F., and Inanc, H., 2013, "Skills at work in Britain: First findings from the Skills and Employment Survey 2012," London.

Fischer, A., Greiff, S., and Funke, J., 2012, "The Process of Solving Complex Problems," Journal of Problem Solving, 4(1), pp. 19–42.

Green, D. M., and Swets, J. A., 1966, Signal Detection Theory and Psychophysics, Oxford, England: John Wiley.

Hwang, W.Y. and Hu, S.S., 2013, "Analysis of peer learning behaviors using multiple representations in virtual reality and their impacts on geometry problem solving," Computers & Education, 62, pp.308-319.

Jin, Y. and Lee, S., 2019, "Designing in virtual reality: a comparison of problem-solving styles between desktop and VR environments," Digital Creativity, 30(2), pp.107-126.

Jonassen, D., Strobel, J., and Lee, C.B., 2006, "Everyday Problem Solving in Engineering: Lessons for Engineering Educators," J. of Engineering Education, 95(2), pp.139–151.

Klaib, A.F., Alsrehin, N.O., Melhem, W.Y., Bashtawi, H.O. and Magableh, A.A., 2020, "Eye tracking algorithms, techniques, tools, and applications with an emphasis on machine learning and internet of things Technologies," Expert Systems with Applications, 166(15), p.114037.

Lee, J. Y., Donkers, J., Jarodzka, H., and van Merriënboer, J. J. G., 2019, "How prior knowledge affects problem-solving performance in a medical simulation game: Using game-logs and eye-tracking," Computers in Human Behavior, 99, pp. 268–277.

Lodge, D. and Teodor, G., 2011. "Training and Assessment of Technical Skills and Competency in Cardiac Surgery," European Journal of Cardio-Thoracic Surgery 39 (3): 287–93.

Macedo, R., Correia, N., and Romão, T., 2019, "Paralympic VR Game: Immersive Game Using Virtual Reality and Video," Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, New York, NY, pp. 1–6.

Man, D. W. K., Poon, W. S., and Lam, C., 2013,"The effectiveness of artificial intelligent 3-D virtual reality vocational problem-solving training in enhancing employment opportunities for people with traumatic brain injury," Brain Injury, 27(9), pp. 1016–1025.

Martinez, M. E., 1998, "What Is Problem Solving?," Phi Delta Kappan, 79(8), pp. 605–609.

Milella, F., 2015, "Problem-solving by immersive virtual reality: Towards a more efficient product emergence process in automotive," Journal of Multidisciplinary Engineering Science and Technology (JMEST), 2(4), pp.860-867.

Mourtos, N. J., DeJong-Okamoto, N., and Rhee, J., 2004, "Open-ended problem-solving skills in thermal-fluids engineering," Australasian Journal of Engineering Education, 8(2), pp. 189-200.

OECD, 2013, "Time for the U.S. to Reskill?: What the Survey of Adult Skills Says," OECD Skills Studies, OECD Publishing, Paris.

Poole, A., and Ball, L., 2006, "Eye tracking in human-computer interaction and usability research: Current status and future prospects," Encyclopedia of Human Computer Interaction, pp. 211–219.

Stanislaw, H. and Todorov, N., 1999, "Calculation of signal detection theory measures," Behavior Research Methods, Instruments, Computers, 31(1), pp. 137–149.

Tang, Y., Shetty, S., and Chen, X., 2011, "Empowering students with engineering literacy and problem-solving through interactive virtual reality games," 2nd International IEEE Consumer Electronic Society Games Innovation Conference, ICE-GIC 2010, Hong Kong, China, pp. 1–6.

Tsai, M.-J., Hou, H.-T., Lai, M.-L., Liu, W.-Y., and Yang, F.-Y., 2012, "Visual attention for solving multiple-choice science problem: An eye-tracking analysis," Computers & Education, 58(1), pp. 375–385.

Wang, P., Wu, P., Wang, J., Chi, H. L., and Wang, X. Y., 2018, "A Critical Review of the Use of Virtual Reality in Construction Engineering Education and Training," International Journal of Environmental Research and Public Health 15 (6).

Wang, Y. and Chiew V., 2010. “On the Cognitive Process of Human Problem Solving.” Cognitive Systems Research, 11 (1): 81–92.

Ward, T. B., 2012, Handbook of Organizational Creativity, Academic Press, San Diego, Chapter 8 - Problem Solving.

Yang, H., Kumara, S., Bukkapatnam, S. T. S., and Tsung, F., 2019, “The internet of things for smart manufacturing: A review,” IISE Transactions, 51(11), pp. 1190–1216.

Zagermann, J., Ulrike, P., and Harald, R., 2018, “Studying Eye Movements as a Basis for Measuring Cognitive Load,” Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, New York, NY, pp. 1-6.

Zhao, R., Aqlan, F., Elliott, L. J., and Lum, H. C., 2019, “Developing a virtual reality game for manufacturing education,” Proceedings of 14th International Conference on the Foundations of Digital Games (FDG). San Luis Obispo, CA, USA.

Figure Captions List

Figure 1	Example stations in the virtual manufacturing system: (a) component selection station; (b) base station; (c) tire and rim station; (d) roof station
Figure 2	Proposed Research Framework
Figure 3	Derivation of sensitivity index d'. HR and HR' are symmetric with respect to $x = \frac{d'}{2}$; FAR and FAR' are symmetric with respect to $x = 0$.
Figure 4	A simple selection problem
Figure 5	Experimental setup (left) and a student running the VR simulation (right)
Figure 6	Heatmap of sensitivity index d'.
Figure 7	Heatmap of problem-solving skills measure ρ'.
Figure 8	The red spider charts of VR-based composite indices computed from 5 scores, including car quality score, station switch score, scores of car weight, price, and cycle time: (a) Index of SME is 85.5951; (b) An example of high index which is 59.7093; (c) An example of low index which is 11.9053.

Table Caption List