Measuring problem-solving skills with virtual reality

VR simulation, eye tracking used to study metacognitive abilities in design and manufacturing

By Faisal Aqlan, Lisa Jo Elliott and Richard Zhao

Problem-solving is an iterative process that requires brainstorming, analysis of the problem, development and testing of solutions. It relies on understanding what is known and what is unknown about the problem. That knowledge of the knowns and unknowns is called metacognition.

Today's engineers must understand their own metacognition and that of other team members to derive the best solutions for engineering problems given the different constraints. Engineers working in design and manufacturing fields confront challenges due to a lack of important metacognitive understanding of their own and their team's problem-solving skills. This research suggests measuring metacognition within teams by using manufacturing simulations with virtual reality and eye tracking.

Engineering education must transform students interested in engineering into professionals who have the knowledge, skills and abilities to create reliable and innovative products. Teaching such knowledge is similar to that in other disciplines yet passing on the skills and abilities to the next generation of engineers is more challenging. Students often come to engineering courses with poor problem-solving skills, an incomplete knowledge of how to best learn and insufficient metacognitive skills.

Through a research project funded by National Science Foundation, the authors studied metacognitive problem-solving in undergraduate engineering students using virtual reality and eye tracking. The project focuses on simulating different types of manufacturing systems. The study was intended to determine if a simulated manufacturing process would help students seek more information and improve their metacognitive skill. To do this, a virtual reality (VR) simulation using a video game environment was created to mimic the physical simulation of building toy cars.

How the simulation works

The simulation activities involve designing and producing toy cars that satisfy customer requirements while minimizing production costs. Customer requirements (constraints) are divided into two main categories: vehicle requirements, including car weight, material and labor costs, color options and size; and functional requirements, including that the driver must be able to get in and out of the vehicle and see where he is going while traveling, the vehicle must be able to travel over ramp conditions, stay on the ramp and cross the finish line fully intact, and the vehicle must remain intact following a drop test.

Producing the toy car involves six job descriptions seen in Figure 1: customer, design engineer, manufacturing engineer, sourcing engineer, quality engineer and supplier.

The goal was to minimize the total cost of producing the car while satisfying the customer's requirements. Hence, there were four main functions: design, sourcing, manufacturing and inspection. To assess the metacognitive problem-solving skills during the simulations, the research measured group effectiveness and metacognitive awareness. An example of customer requirements for an order is found in Figure 2.

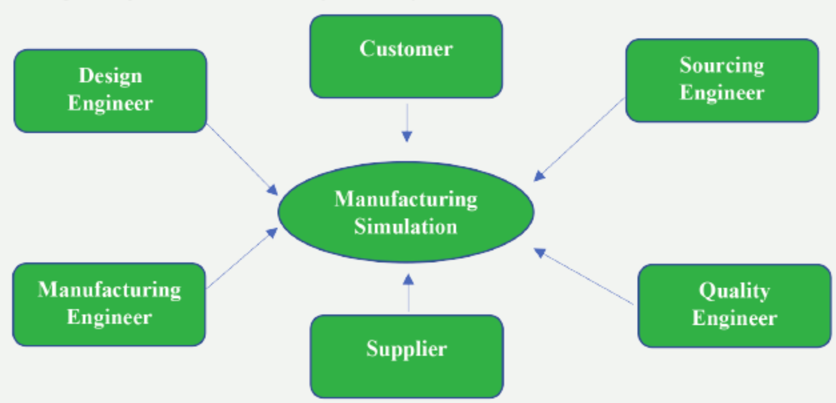
Physical simulation of toy car assembly

The physical or traditional simulation consists of the assembly of a toy car made of plastic bricks. The simulation is conduct-

FIGURE 1

Establishing the roles

Producing the toy car involves the six job descriptions listed here.



A look at the simulation kit, above, and workstation layout, inset, used in the physical simulation assembly.

FIGURE 2

Meeting the customers' needs

An example of the customer requirements for an order of toy cars.

Car requirements	Functional requirements
(a) vehicle weight between 20 and 40 grams (b) material cost ≤ $10 (c) number of individual components ≤ 2 (d) vehicle must fit completely within the design footprint "parking space" (e) number of different types of Lego blocks ≥ 10 (g) number of different colors for Lego blocks = 5 (exclude driver and wind shield) (h) vehicle must have four tires (with axles), wind shield, driver, steering wheel, and roof	(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

ed by students individually and in groups. The accompanying photos show a layout of the assembly line as well as the simulation kit.

To improve the kitting process, the students developed an optimization model for the assembly of the toy cars. The model takes into consideration the customer requirements of the vehicle to provide the best layout for assembling the toy car while minimizing the production cost. The objective function is:

$$max \quad \sum_{k=1}^{K} Z_k \left(p_k - (\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{r=1}^{R} c_{ij}\, x_{ijrk} + h \sum_{i'=1}^{I'}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{L} x_{ijrk}\ (\ t_{ii'} + t_{il})) \right)$$

In this, Zk is a binary decision variable (if the order k is produced, then Zk = 1, otherwise Zk = 0); pk is the selling price of order k; Cij is the car component i of size j; Xijrk represents car component i of size j and color r to be used in order k; h is the operator cost per time unit; Tii' is the time to assemble component i with another component i'; and Til is the time to move part i from location l.

Several constraints were considered. For example, the weight constraint is set as:

$$W_{min} \leq Z_k \left(\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{r=1}^{R} w_{ij}\, x_{ijrk} \right) \leq W_{max}\ , \forall k \in K$$

In this, Wij is the weight of component i of size j and wmin and wmax are the minimum and maximum weight limits, respectively. Sample results for the assembly time for four different toy car options and two production layouts are included in Figure 3.

An example of the requirements of a given customer order is shown in Figure 4.

Virtual reality simulation model

VR technology brings immersion to the next level by allowing a user to become fully immersed in a different world. The popularity of VR in recent years has been helped by the introduction of lightweight, affordable headsets. In this study, the researchers produced a VR simulation environment that mimics the physical simulation and uses eye tracking to measure the metacognitive skills of students while they solve the problem of toy car design and production.

The VR simulation was built in the Unity game engine with the HTC Vive VR headset, wireless controller and base stations for motion tracking. The headset was custom fitted with Tobii eye-tracking technology, allowing the system to identify the coordinates and objects a user was looking at during any given time in the simulation. A "user" refers to a student or a participant wearing the headset using the VR simulation.

In the simulation, the user saw through the headset a virtual environment consisting of a series of stations and could then interact with the objects in the virtual environment, such as picking up a virtual plastic piece using the wireless controller.

Since VR is a relatively new technology, a user's previous experience with such an environment must be taken into account. In this simulation, users were first presented with audio instructions on how to interact with the virtual environment. Once the user was comfortable, a button was pressed to start the actual manufacturing process and a timer started counting.

An example of a toy car that was produced in the physical and VR simulations is shown in the photos on page 44. The researchers recruited 12 male engineering undergraduates with an average age of 19 who reported they had taken an average of 70 credit hours. Of the 12, 11 contributed eye tracking in the VR simulation.

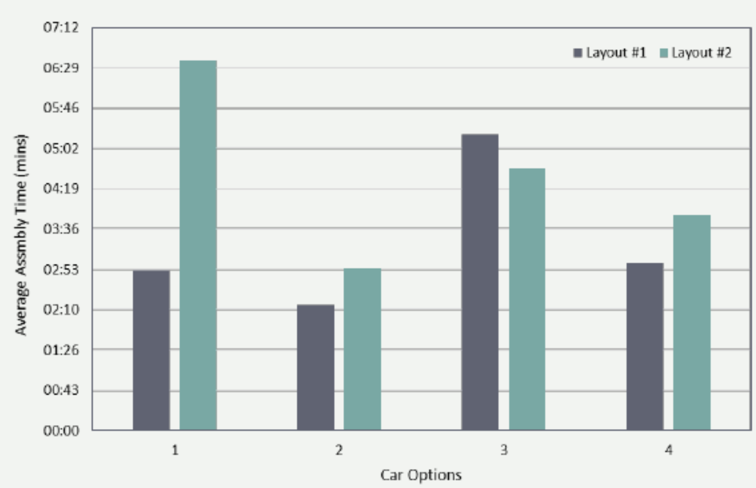
Six participants successfully completed the task within the allocated 20 minutes. The average assembly time was 8 minutes, 41 seconds to complete the work on five stations for toy car option 4. This is also comparable to the physical simulation assembly time. Figure 5 shows a breakdown of the average time in seconds spent by participants at each station.

Among the assembly stations, participants spent the most time on the sides station. Recorded videos of the sessions showed that this station was where the participants started making corrections to their previous choices in the plastic pieces as they started to examine more closely whether they were able to meet the requirements.

FIGURE 3

Assembly speed

Car toy assembly time for two different production layouts.



Measuring metacognitive skills using eye tracking

The researchers incorporated the raw eye tracking data as a measure of metacognition into a model of conflict and error to predict what types of experiences are most beneficial when training metacognitive skills. The raw eye tracking data analysis uses signal detection theory (SDT) to differentiate stimuli and quantify a student's performance. Initially, the expectation was that the student would survey the car parts available and focus on one or two options to manufacture the car.

Through comparing the student data to the expert data, researchers obtained a more accurate estimate of what the student was considering and how metacognition is developed by sensitivity (an observer's ability to discriminate stimuli) and response bias (an observer's standards for producing different behavioral responses). It measured the viewing of items they attended to, the amount of time and the order in which they attended to them and their choice of attention to each.

Information from the eye-tracker collected two basic measures: gaze fixation and saccade, a rapid movement of the eye between fixation points. Previous literature established these as sufficient measures of attention and information processing related to learning in knowledge change and metacognition. Fixation measured the amount of attention in terms of location (area of interest, or AOI) and in terms of time, and the

FIGURE 4

Order specification

An example of the requirements of a given customer order.

Component	Size	Color	Quantity
Brick	1x1	Yellow	1
Brick	1x3	Green	1
Plate	2x6	White	1
Steering Wheel	1x2	Black and White	1
Wind Shield	3x6	colorless	1
Tires	Small size	Black	4
Axle	Small One Side	Black	4
Rim	Small	White	4

FIGURE 5

Measuring time

The average time in seconds participants spent at each station.

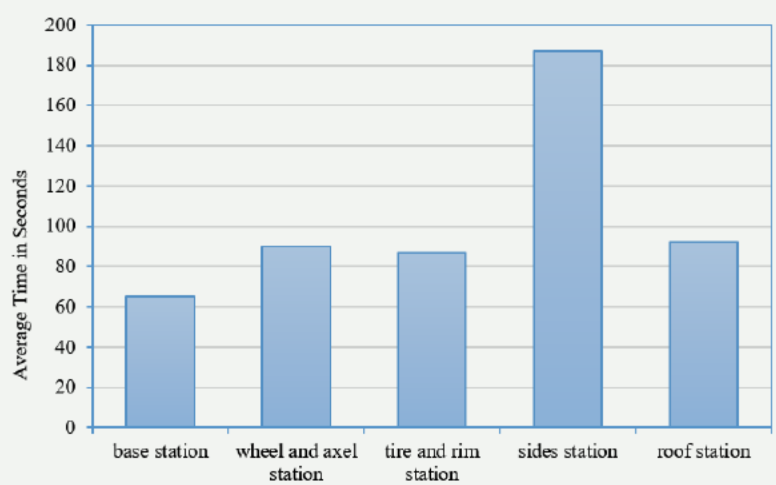


FIGURE 6

Measures of eye tracking

Spatial and temporal information gathered from the eye tracker.

	Spatial Measure	Temporal Measure
Fixation (attention- correct item)	Gaze Point- as an X,Y coordinate as the center of the optimum Area of Interest/AOI (attention to correct item)	Time stamp of gaze point – gathered each 16.7 millisecond
Saccade (order of processing, correct length of time)	Order of gaze points- what the participant looks at first, second, third, and so forth	Loiter of fixation- the amougont of time the participant gazes before the next saccade (attention- correct length of time)

Participants use virtual reality headsets, top, to complete the assembly process; an example of what they were seeing is at bottom.

saccade measured the length of time for which items on the screen are attended to. Within these two measures, there was spatial and temporal information as seen in Figure 6.

To quantify expertise, comparing the items the participants looked at to the items the expert looks at provided three categories: hits, misses and false alarms. For each, these definitions were used.

- A hit is when the expert used this item and the number of times that the expert used the item. For example, if the expert used a particular piece five times and a participant used the same piece seven times, only five would be hits.
- A false alarm is when the expert did not use this item but the participant did; this included items the participant used more than the expert. In the example for hits, the two pieces the participant used that were in excess count in this category.
- A miss is when the expert did not use an item the participant used.

Traditionally, correct rejection is also included in signal detection theory. This could be calculated by taking the total number of items available and the number of times they could be chosen. In this case, the items could be chosen infinitely. This calculation was not useful in determining understanding and accuracy in the task.

The 11 participants who contributed to eye tracking looked at an average of 2,573 items with a standard deviation of 1,625. The fewest number of items one looked at was 609 items; the most was 5,399. This compared to one subject matter expert who completed the simulation and looked at 664 items.

FIGURE 7

Measuring hits, false alarms

A chart shows the number of hits on the X axis and the number of false alarms on the Y axis. The size of the dots correspond to the participant's score, with better scores having larger dots.

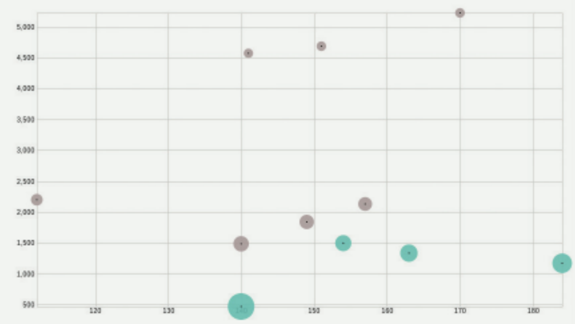
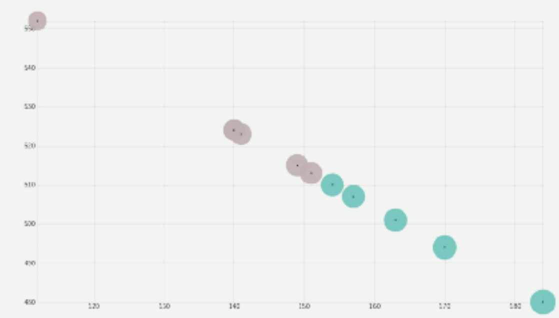


FIGURE 8

Measuring hits, misses

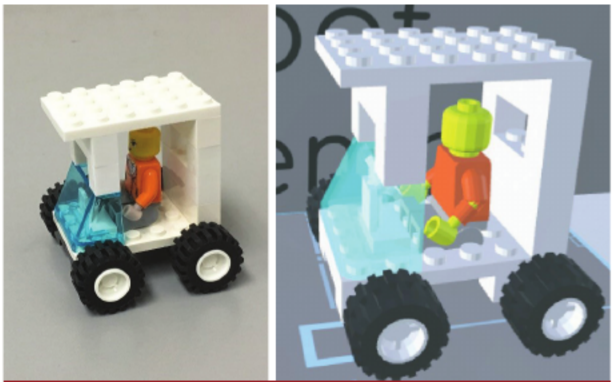
A chart shows the number of hits on the X axis and the number of misses on the Y axis. The size of the dots correspond to the participant's score, with better scores having larger dots.



Participants displayed an average hit rate of 151 items with a standard deviation of 19. They had an overall miss average of 513 items with a standard deviation of 19. The false alarm rate was 2,422 items with a standard deviation of 1,624.

Figure 7 shows the number of hits on the X axis and the number of false alarms on the Y axis. The size of the dots corresponds to the participant's d` score, with better scores having larger dots. A hit rate of 664 and a false alarm rate of 0 would correspond to a perfect score of 1.0. Most participants' score was at 0.03.

Only two participants had a hit rate above 166, but with false alarm rates in the thousands, their score was low. The



Examples of the assembled toy car from the physical simulation, at left, and the virtual simulation, at right.

Va. Tech plans virtual construction sites to train engineers

The hands-on nature of construction engineering makes it difficult to train new workers in a cost-effective manner, as doing so requires time, labor, costly materials and safety risks. But a proposed virtual reality learning platform would give students a chance to learn "on the jobs."

Abiola Akanmu, an assistant professor in the Myers-Lawson School of Construction at Virginia Tech, and an interdisciplinary team of Virginia Tech researchers is creating virtual construction scenarios by using augmented reality and holograms. They recently received a National Science Foundation grant to fund their efforts.

"Educators and commercial firms are both looking for ways to solve the hands-on problem," Akanmu said. "And whatever you can do in the physical world, we'll be able to do in the virtual world."

The research team will use laser sensor technology to map construction sites and create a "smart" virtual reality job site. From there, they create scenarios at the virtual site through AR and holograms that allow construction engineering and management students to experiment using Microsoft HoloLens. The hands-on training comes without concerns about errors that could inflict damage or injury and expend time, labor and material costs. Akanmu also envisions construction engineering students will learn safe ergonomics practices.

The research will help educators tailor student experiences through feedback from focus groups and experienced contractors in the construction industry.

"Students will learn not just the technical skills, but also the decision-making skills that firms seek in their employees to make them more effective," said Bolanle Ogunseiju, a doctoral student in the environmental design and planning program.

Source: Virginia Tech Daily, vtnews.vt.edu

top two participants had hit rates of 170 and 184 and scores of 0.16 and 0.30, respectively. This analysis describes how well the participants were able to attend to the necessary items in the simulation and ignore the distracting noise or false alarms.

The two participants scoring above 0.10 were 19- and 20-year-old freshmen engineering students. Their metacognition scores on the State Awareness Questionnaire (SAQ) were at the median score in awareness, planning, self-checking and cognitive strategy.

A second analysis explored if hits and misses described participants' performance more accurately. Again, the average hit rate was the same: 151 items with a standard deviation of 19. The average miss rate was 513 items with the same standard deviation. The d` scores were higher in the second analysis with a range of 0.20 to 0.38. The same individuals who had the two highest scores in this analysis were also those who had the two highest scores in the first analysis (in Figure 7, bottom right).

In this analysis, five participants' d` scores were 0.302 to 0.3833 with the remaining five scoring between 0.2029 and 0.2945. The lowest score in this analysis was 0.302; the highest was 0.3833.

This analysis describes how well the participants were able to discern what should and should not be on the car. Figure 8 shows the number of hits on the X axis and the number of misses on the Y axis. The size of the dots corresponds to the participants' d` score, with better scores having larger dots.

As shown in the figure, the data points are perfectly aligned as the sum of hits and misses must always equal the expert's score and is a ratio of how many times the person correctly assessed an item's inclusion in the car. ❖

Faisal Aqlan, Ph.D., is an assistant professor of industrial engineering at the Behrend College at Pennsylvania State University . He earned his doctorate in industrial and systems engineering from the State University of New York at Binghamton in 2013. He is a member of IISE and has received numerous awards and honors including the IBM Vice President award for innovation excellence.

Lisa Jo Elliott, Ph.D., is an assistant teaching professor of psychology at the Behrend College at Pennsylvania State University. She received her bachelor's degree from the University of Arizona in Tucson, Arizona, and her master's and doctorate degrees in experimental psychology with an emphasis in engineering psychology from New Mexico State University in Las Cruces, New Mexico in 2011.

Richard Zhao, Ph.D., is an assistant professor of computer science and software engineering at the Behrend College of the Pennsylvania State University. He received his bachelor's degree in computer science from the University of Toronto and his master's and doctorate degrees in computing science from the University of Alberta in 2009 and 2015, respectively. His research focuses on artificial intelligence in video games and machine learning and data mining techniques in social media.