

# Evaluating Hazard Awareness in Virtual Reality Collaborative Robotics Using Eye Tracking

Joshua Odeniran & Yee Mey Goh

Loughborough University

---

## SUMMARY

This study evaluates the use of Virtual Reality (VR) as an effective training method of assessing operators' hazard awareness when working with robots in collaborative tasks. The VR environment used in testing uses the UR\_16e robot with three scenarios where a specific robot variable changes with the aim for the operators to notice. Findings indicate that participants recognised two of the hazards (force and speed changes) and were able to explain the adverse effects. Although directional changes were less frequently identified, this study shows that VR training contributed to their understanding of robotic behaviour.

## KEYWORDS

Hazard Awareness, Human Robot Collaboration, Training

---

## Introduction

In traditional industrial environments, tasks were performed using mechanical machinery under the continuous supervision of trained operators, ensuring a smooth and efficient workflow. Operators were responsible for identifying and addressing malfunctions in real time. However, as consumerism increases and the demand for specific products rises, these industrial environments have needed to automate their processes, implementing autonomous and collaborative robots (cobots) to improve production and efficiency. This change requires trained operators to work in closer proximity with these cobots, utilising human problem-solving skills alongside the repetitive and precise capabilities of robots. This technological transformation is commonly referred to as Industry 4.0.

Currently, in developed economies, the market for these cobots has grown exponentially as Industry 4.0 has developed and progressed. The productivity and efficiency of cobots have increased; however, the emergence of human-robot collaboration (HRC) has created new hazards and issues, primarily due to the mental stress associated with working alongside robots (Giallanza et al, 2024). This mental stress can lead to unsafe work behaviours, resulting in injuries and, in certain cases, death (Evangelista et al., 2024). To reduce the number of workplace injuries, suitable training must be utilised to cultivate safe workplace behaviours.

This research evaluates whether VR can serve as an effective platform for training hazard awareness in collaborative robotic environments. The study combines eye tracking data (fixations and saccades) with Situation Awareness Global Assessment Technique (SAGAT) testing to provide a cognitive assessment.

Training operators on how to work safely with robots is an expensive and time-consuming process, as companies rely on an external firm to send a certified instructor to train employees. These instructors face limitations, such as restricted time and resources, which can lead to some operators having a weaker understanding.

However, with the development of VR technology, virtual environments have become increasingly accessible and can effectively replicate a factory workspace. Additionally, a VR training module is more accessible than an external trainer and can accommodate a person's schedule.

To assess the suitability of VR for operator training, this study will address the following key objectives:

- Create a suitable VR training module based on the UR\_16e
- Utilise SAGAT questions, eye tracking and questionnaires, to understand participants' attitude towards robots and how that affects their hazard awareness
- Measure changes in participants' learning level and how their hazard awareness develops.

The remainder of this paper is structured as follows: a review of relevant literature, the research methodology, analysis of results and discussion of findings and future implications.

## Method

### *VR Game Design*

The study used an immersive VR environment, replicating a manufacturing shopfloor, with the UR\_16e robot in the centre. The task and environment were developed using the HTC VIVE Pro Eye and Unity 6000.0.43f1. Traversal of the stage used snap locomotion.

3D models from the Unity Asset store such as shelves, carts and industry machinery were implemented to enhance the factory's realism. The models were then arranged to reflect a typical factory layout, increasing user immersion. This is particularly important, as immersion is fundamental in making people feel present and engaged, as sensory realism, primarily audiovisual, greatly supports the participants' immersion (Potter et al 2022).

Steam VR was used to integrate the HTC VIVE Pro Eye and Unity. The set-up enabled the headset's eye-tracking capabilities to function within the virtual environment, allowing for the collection of gaze and fixation data necessary for the study. Figure 1 depicts the environment created.



Figure 1: VR factory environment

A similar approach (Choi A et al 2020) was taken in the task design. Participants operated in a controlled VR environment, directly interacting with a simulated UR\_16e robot to complete a

simple task. The robot behaviour changed in each of the following three levels, aligning with ISO 15066:2016 hazard categories (i.e. speed, direction, and force).

### ***Sample and Participants***

Participants were recruited from Loughborough University with diverse experience with robots, gender and degree disciplines. Following ethics and risk assessments, participants with vertigo, epilepsy or who were pregnant were excluded from the study. Participants did not have industry experience; this mitigates against prior experience or training skewing the results.

Inclusion criteria:

- Participants must be 18+
- Normal or corrected to normal vision
- Participants must be able to tolerate VR

In this study 16 participants were recruited in total with the experiments completed in the period of 30-45 minutes. Table 1 highlights their distributions of age, gender and proficiency with robots.

Table 1: Distribution of participants

<b>Participants</b>	<b>Gender</b>	<b>Age</b>	<b>Proficiency with Robots</b>
A	F	22	Not Proficient
B	M	21	Not Proficient
C	M	22	Not Proficient
D	M	21	Not Proficient
E	M	22	Not Proficient
F	M	22	Proficient
G	F	21	Not Proficient
H	M	21	Proficient
I	M	24	Proficient
J	F	22	Not Proficient
K	F	23	Not Proficient
L	M	22	Proficient
M	M	23	Not Proficient
N	F	21	Proficient
O	M	19	Proficient
P	M	22	Proficient

Ethical approval was given by Loughborough University's ethics committee (21489).

### ***Procedure***

As shown in Figure 2, participants were required to complete a tutorial to familiarise with the VR task environment and three types of hazards working with the robot; Level 1: Speed change, Level 2: Directional change and Level 3: Force applied. These variables were selected due to the safety requirements for collaborative robots outlined in ISO/TS 15066:2016.

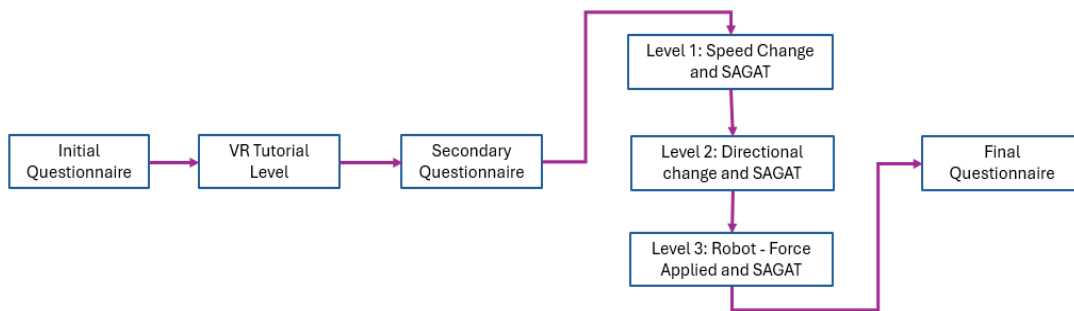


Figure 2: Experiment workflow

To accurately collect data, participants invited to take place in the study were provided with food and water to aid their comfort whilst the requirements of the study were once again explained to them. As stress can increase cognitive load it was imperative that the participants remained as comfortable as possible. Figure 2 highlights the workflow; however, a 15-minute break was offered to the participants after each section utilising VR, i.e., training, Level 1, Level 2 and Level 3. This meant the overall study could potentially take 1 hour.

### Data collection

During the study, their eye tracking data (fixation) was recorded and exported to an Excel file. Participants were asked questions during the levels and filled in a questionnaire at the end to evaluate what they had learned.

The primary aim of this training module is to assess hazard awareness through eye tracking (fixations and gaze duration) and SAGAT responses. Therefore, task completion speed was not used to evaluate the participant's performance, as the objective was to observe cognitive engagement as behaviours rather than the speed of task completion.

The participants' fixations on specific areas of interest (AOIs), such as the robot's arm and the block placement zones were recorded. While participants completed the task, SAGAT questions were used to test the users' situational awareness level. Participants were asked to complete a pre-/post-questionnaire to gauge their experience with VR and robots and respond to task-related questions assessing their level 3 situational awareness. Cognitive task analysis was then utilised to link the fixation data with the SAGAT questions and pre-/and post-task questionnaires, establishing a connection between participants' fixations and their behaviours (Kushniruk et al 2004).

SAGAT testing was chosen as it can gauge level 1-3 situational awareness. These questions were embedded at intervals across the task, whilst participants were not necessarily focused on the task, ensuring immersion was not broken. The questions were designed to measure user perception, comprehension, and application in real time. User responses were categorised into different reaction and learning levels based on the Kirkpatrick model. (Evangelista et al, 2024). An example of the SAGAT questions asked are as follows:

- Based on our tutorial and conversation what may go wrong?
- What areas will you be looking at to know if there is a problem or not?
- What will you do when problems arise?
- What will happen if this problem goes unchecked?

### Results

A structured analysis of all the participants' eye tracking, questionnaire and SAGAT results was collated and analysed to determine if VR was a suitable way of measuring hazard awareness within

participants. 16 participants were recruited, and data sets were collected for each of them. The data collected was split into 2 groups a robot proficient group (7 participants) and a non-robot proficient group (9 participants).

#### a) Fixation Behaviours

T tests were used to analyse the results of the tests to prove the null hypothesis **H<sub>0</sub>: There is no significant difference in fixation data between participants with robot proficiency and those without.** This hypothesis was proposed to establish there is not a large difference in fixation data between robot proficient and non-proficient users. In each of the levels there were no significant differences in fixations between the 2 groups, as Level 1: Speed change ( $p = 0.625$ ), Level 2: Directional change ( $p = 0.98$ ) and Level 3: Robot force applied ( $p = 0.296$ ).

#### b) Hazard Detection

Although there was no significant change in fixation data between both groups of participants the use of a questionnaire showed strong overall hazard detection across participants. Overall, 81% of participants were able to identify two or more of the variables which changed in the robot's behaviour, regardless of their prior proficiency with robots. The two correctly identified variables were Level 1: Speed changes (56%) and Level 3: Robot force applied (81%), whilst Level 2: Directional changes (31%) were less frequently reported.

#### c) Learning Outcomes & Situation Awareness

Within the study, there was a trend where, Level 2: Directional change showed slightly lower fixation times than in Level 1 for the robot proficient group. As participants became more familiar with the task this can lead to a premature disengaging with the robot's motion. This may help explain why directional change was the least reported variable. SAGAT questions prompted participants to anticipate potential failures and show that participants were focused on task performance rather than robot behaviour, suggesting limited awareness of the robot and how it contributes to potential hazards.

### Discussion

The study's findings largely supported the hypothesis, in which there was no statistically significant difference in fixation data between participants with robot proficiency and those without. The SAGAT questions and questionnaires support these findings; however, trends emerged in the two groups, proficient (PR) and non-proficient (NPR).

#### a. Level 1

In level 1, the robot's movement speed varied. At this level, the total fixation time is the highest within the PR group at **mean = 131.49s** and the second highest within the NPR group at **mean = 114.85s**. These high fixation times suggest heightened vigilance following the tutorial. (Endsley, 1995) model of situational awareness, states that prolonged fixations correlate to level 1 situational awareness, which reflects the user's perception of their environment. Participants would encounter unpredictability, which in this context would be the robot's movements; this prompts participants to spend more time fixated on the robot to familiarise themselves with it during the tasks.

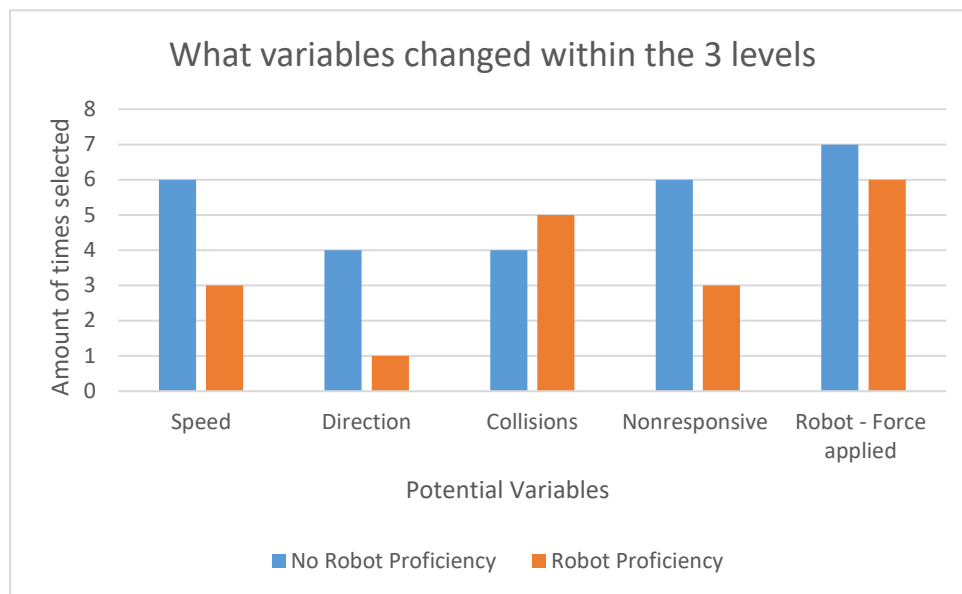


Figure 3: Questionnaire results in split groups

Although Level 1 had the highest mean fixation time, speed was the joint second-most selected option that changed within the level as seen in Figure 3. This demonstrates that while a participant visually engages with the robot, it contributes to their Level 1 situational awareness. Prior studies suggest a correlation between fixations and active recall of memory ((Bernhard Fehlmann et al., 2020))

This explains why, although it was the first variable that was changed at the beginning of the study, most participants remembered the variable, as fixations more often increased free recall performance. This was further supported by the fact that 65% of participants recalled speed as a changing variable. Interestingly, the NPR participants recalled the change more frequently (67%) than the PR participants (50%), even though the mean fixation time was lower than that of the PR participants. This implies that prior familiarity with robots did not guarantee better recall.

#### b. Level 2

Direction was the variable that changed within this level, Figure 3 shows that direction was the least frequently picked option. In this level, the mean fixation time dropped significantly within the PR group to **116.48 seconds**; however, in the NPR, the mean total fixation time increased to **117.07 seconds**. This reduction in fixation time was as reflected in the results. However, direction was the least selected option in the PR group, only 1 out of 7 participants selected it, whilst in the NPR group, 4 out of 9 participants selected it. These results suggest 2 separate sub-trends; one found within the PR participants and the other within the NPR group.

Within the PR participants group, whilst there was a significant drop in the mean fixation time from the previous level, it was still within 1 second of the mean of the NPR group, which does not explain the drastic change where the NPR participants were able to recognise the changing variable, and they were not. However, through the SAGAT questions, there was a noticeable trend within both groups. In the PR group, through observation, the participants started to traverse the space faster in VR, completing the positioning of the cubes faster. This suggests that after the first level, users got more comfortable with the task, indicating they felt less need to pay as much attention or put as much effort into it. Prior research indicates that cognitive biases, including unwarranted

confidence in one's judgments, can lead individuals to underestimate the probability of failure in complex tasks (Tversky et al 1974). This may contribute to reduced vigilance during task execution. This mentality explains the drop in fixation for the PR group, which reflects a breakdown between Level 1 and Level 2 situational awareness. This explains the negative trend where fixation time reduced the more time spent completing the level

In Level 2, participants were asked, "Based on our tutorial and conversation, what may go wrong?" However, both groups gave answers stipulating that they may misalign the blocks or incorrectly orient them, resulting in robotic failure. This highlights an underlying trend in which users were more aware of their actions, causing mistakes rather than robots. This was later supported by question 9 of the SAGAT questions, where users were asked, "Did seeing any of these issues change how you view robots?". Participants began to realise, as one noted, that "robots are not infallible" and require "someone overseeing it". This self-reflection implies that participants assumed confidence in the task and did not weigh all the issues that may occur effectively. This assumed confidence in their knowledge of robots, particularly in the PR group, created a lack of awareness, translating to a poorer detection and recognition of the issue.

### c. Level 3

In Level 3 the robot was programmed to have a changing force output mid task. Unlike in other levels the visual representation of the change was less drastic and only occurred once. This variable was designed to test Level 2 and Level 3 of situational awareness (Endsley 1995), requiring participants, to perceive the change and comprehend the potential impact on task performance.

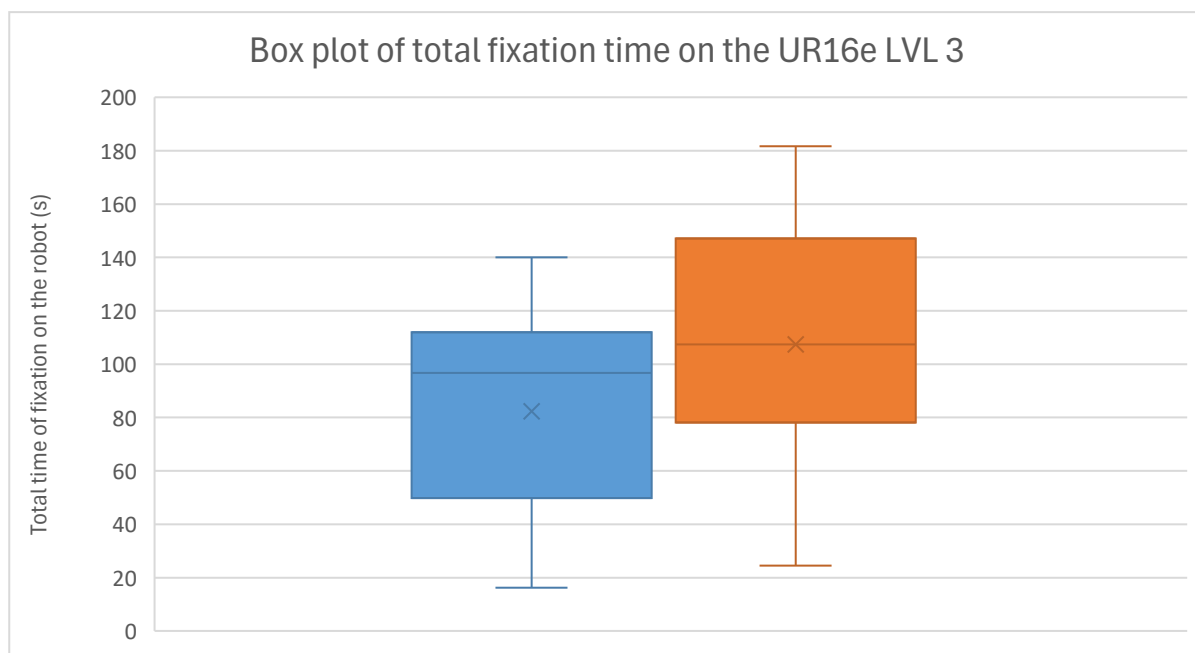


Figure 4: Fixation time boxplot in level 3

As seen in Figure 4 both groups, PR and NPR, the mean fixation time dropped considerably, in contrast it was the most recognised variable change in both groups, this suggests improved change recognition. This was contrasted to level 2 where the PR group underperformed in variable recognition.

In this round of SAGAT testing, participants were asked "Based on our tutorial and conversation, what may go wrong?" At this point of the study there was a noticeable shift in participants answers

from Level 1 to Level 3. In level 3 participants were more aware of the possibility of the robot to make a mistake such as “the robot putting the block in the wrong place.” One possible explanation for this change is task learning and increased cognitive adaptation by the third level. As participants are more familiar with the environment, there is a reduction in cognitive load associated with spatial navigation (Sweller, 1988). Suggesting that participants are more familiar with the environment and task, allowing them to complete the task quicker and more efficiently. This reduced time suggests that participants are more efficient with their fixations. This could explain why force change was the most recognised variable change. This translated into participants realising that “robots are not infallible”.

#### d. Other noticeable trends

In Figure 3 the joint second most picked variables were both collision and the robot being non-responsive. Although these were not variables which were changed in the level, both of which were results of the variable changes.

The non-responsiveness of the robot was due to two different factors in the study. Primarily in level 2 participants who were not paying attention to the robot’s movement, those with low fixation times, were pressing the robot’s activation button whilst the robot was already carrying out a previous task. This resulted in the button breaking and the robot being unable to carry out subsequent tasks from then onwards. Participants were recorded being confused, questioning the robot’s actions and assuming there was something wrong with the robot, all while unaware of their mistake. Although not directly part of the changed variables this correlates to a break down in participants level 1 situation awareness, as they were unable to perceive the robot’s key element, which is it being in motion. This led to a break down in their level 2 situational awareness as participants were unable to explain why this issue happened.

### Conclusion

In conclusion this study investigated the development of hazard awareness in a VR setting during HRC suggesting that VR has strong potential to support the development of situational awareness when working with robotic systems. By leveraging data collection techniques from eye tracking, SAGAT and questionnaires, participants were assessed on their ability to detect changes in robot behaviour across three tasks, enabling an evaluation of their hazard awareness in the VR environment. Future collaborative work with industry would be necessary to develop this technology, ensuring the tasks reflect current processes and aligning the training module with ISO standards.

### References

- Bernhard Fehlmann, Coynel, D., Schicktanz, N., Milnik, A., Gschwind, L., Hofmann, P., Papassotiropoulos, A. and J.-F, D. (2020). Visual Exploration at Higher Fixation Frequency Increases Subsequent Memory Recall. *Cerebral cortex communications*, 1(1). doi:<https://doi.org/10.1093/texcom/tgaa032>.
- Choi, M., Ahn, S. and Seo, J. (2020). VR-Based investigation of forklift operator situation awareness for preventing collision accidents. *Accident Analysis & Prevention*, 136, p.105404. doi:<https://doi.org/10.1016/j.aap.2019.105404>.
- Endsley, M.R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, [online] 37(1), pp.32–64. doi:<https://doi.org/10.1518/001872095779049543>.
- Evangelista, A., Giliberti, C Giglio, V.D., Manghisi, V.M., Raffaele Mariconte and Uva, A.E. (2024). From knowledge to action: Assessing the effectiveness of immersive virtual reality

training on safety behaviors in confined spaces using the Kirkpatrick model. *Safety Science*, 181, pp.106693–106693. doi:<https://doi.org/10.1016/j.ssci.2024.106693>.

Giada La Scalia, Giallanza, A., Manuela, C. and Micale, R. (2024). Occupational health and safety issues in human-robot collaboration: State of the art and open challenges. *Safety Science*, 169, pp.106313–106313. doi:<https://doi.org/10.1016/j.ssci.2023.106313>.

International Organization for Standardization. (2016) Robots and robotic devices – Collaborative robots (ISO 15066:2016). Geneva: ISO. doi:  
<https://www.iso.org/obp/ui/en/#iso:std:iso:ts:15066:ed-1:v1:en>

Kushniruk, A.W. and Patel, V.L. (2004). Cognitive and usability engineering methods for the evaluation of clinical information systems. *Journal of Biomedical Informatics*, [online] 37(1), pp.56–76. doi:<https://doi.org/10.1016/j.jbi.2004.01.003>.

Potter, T., Zoran Cvetković and Enzo De Sena (2022). On the Relative Importance of Visual and Spatial Audio Rendering on VR Immersion. *Frontiers in signal processing*, 2. doi:<https://doi.org/10.3389/frsip.2022.904866>.

Sweller, J. (1988). Cognitive Load during Problem Solving: Effects on Learning. *Cognitive Science*, [online] 12(2), pp.257–285. doi:[https://doi.org/10.1207/s15516709cog1202\\_4](https://doi.org/10.1207/s15516709cog1202_4).

Tversky, A. and Kahneman, D. (1974) ‘Judgment under uncertainty: Heuristics and biases’, *Science*, 185(4157), pp. 1124–1131.<https://doi.org/10.1126/science.185.4157.1124>