# Measuring the impact of Augmented Reality warning systems on onsite construction workers using object detection and eye-tracking

Shaoze Wu, Lei Hou*, Haosen Chen, Guomin (Kevin) Zhang
School of Engineering, RMIT University, Melbourne 3000, Victoria, Australia
*Corresponding Authors: lei.hou@rmit.edu.au

**Abstract.** The failure to identify unsafe conditions has been considered a major contributing factor to construction accidents. Augmented Reality (AR) goggles have emerged as promising devices for warning onsite workers of potential hazards using dynamic visual stimuli. However, there is a lack of methods available to quantify the impact of AR warnings on construction workers, hindering the evaluation, improvement, and promotion of warning systems. To address this gap, a method to measure the impact was developed in this study. Six metrics were designed to quantify the performance of system users under the impact of AR warnings. Next, a module integrating object detection and eye-tracking was proposed to obtain the metrics. Finally, a case study was conducted to validate the feasibility of the proposed method. The results show that the proposed method can precisely capture the subtle eye movements of users and successfully quantify the effectiveness of warning systems.

## 1. Introduction

The construction industry is considered one of the most dangerous sectors with a significant number of accidents occurring every year. A significant contributing factor to construction accidents is the failure to promptly identify hazards (Abdelhamid and Everett, 2000). Studies have demonstrated that showing workers potential hazards on a construction site can effectively reduce accidents (Teizer et al., 2010). With the development of visualisation technology, wearable Augmented Reality (AR) has been introduced in the construction safety area to alert onsite workers to potential threats through visual prompts (Wu et al., 2021). AR devices can visualise real-time hazard information, such as hazardous areas, as holograms in the real-world environment, allowing workers to remain focused on their tasks while also being aware of surrounding hazards with a quick glance. For example, Kim et al. (2017) developed a wearable AR system that uses arrows and text to indicate the direction and distance of the nearest vehicle to onsite workers. Wu et al. (2022) developed a visual warning system to identify construction hazards through onsite cameras and show the hazard information (i.e., hazard sources, hazard types, and hazardous areas) to onsite workers using AR goggles. These studies have demonstrated the technical feasibility of using AR technology to improve construction safety. However, research on AR warning systems is still in its early stages, and further development is required to effectively implement such systems in practice. Currently, a significant challenge is the lack of a suitable method to measure the impact of AR alerts on onsite construction workers. The absence of this method makes it difficult to evaluate and enhance AR warning systems, impeding the development of wearable AR for enhancing construction safety.

Eye-tracking is a valuable tool for evaluating the usability of Human-Computer Interfaces (HCIs) and has been extensively used in high-risk industries such as aviation, maritime, and construction (Martinez-Marquez et al., 2021). By providing detailed data on eye movements, eye-tracking enables researchers to quantitatively evaluate the strengths and weaknesses of a system. For instance, Hareide and Ostnes (2016) demonstrated the effectiveness of a simulator for navigation training by analysing visual focus distribution data. Li et al. (2019) used eye-tracking to analyse the interaction between pilots and flight deck interfaces. Chew et al. (2018)

evaluated the impact of in-vehicle visual support on construction equipment operators using eye-tracking glasses. Despite the proven benefits of eye-tracking in evaluating digital systems, applying eye-tracking techniques to measure the effectiveness of wearable AR systems has not received sufficient attention. Therefore, this study aims to develop a quantitative approach to measure the impact of AR warnings on onsite construction workers using eye-tracking technology. The article is organised as follows: Section 2 reviews the onsite safety warning systems and corresponding evaluation methods, as well as the application of eye-tracking in the construction safety area. Section 3 introduces the research methodology. Section 4 presents and discusses the results of a case study. Finally, the paper is concluded in Section 5.

## 2. Literature Review

### 2.1 Safety Warning Systems and Evaluation Methods

With the incorporation of more technology, safety warning systems offer multimodality to help onsite workers obtain hazard information in a natural and intelligent manner. As a result, evaluating the usability of such warning systems requires more sensitive and accurate methods. Teizer (2015) introduced Radio Frequency (RF) to detect the proximity between workers and equipment, and he used beepers to draw workers' attention to potential safety hazards. Due to the large amount of noise in the construction site, using auditory signals alone may not be effective in delivering hazard information to workers (Teizer et al., 2010). Kanan et al. (2018) added vibrators to their proximity warning system to provide tactile warnings. In these studies, sound and vibration were used to help workers know whether they were in danger, so the experiments focused on the accuracy of hazard identification rather than the impact of warnings on workers. Huang et al. (2021) coded the vibration frequency to enrich the information it can convey. They designed three vibration patterns with different intervals for a wristband and tested the warning effectiveness. Experiment results showed that most workers could distinguish the different vibration patterns. The workers believe that a shorter interval represents a higher degree of urgency. Sakhakarmi et al. (2021) developed a haptic-based wearable hazard communication system and experimentally verified the feasibility of the system. The system was configured with 10 vibration motors at different locations on workers' backs to indicate eight relative directions of hazards, three hazard levels, and two hazard types using different combinations of vibrations. In the experiment, the eyes and ears of their experiment participants were covered to simulate obstructed vision and loud noise, and the participants were asked to move to a safe area after detecting the surrounding hazards. The experiment validates the feasibility of the system, but it hardly reflects the usability of the system. The participants were fully focused on feeling the vibrations and were not involve in any construction activities, which is rare in real-life use cases.

The rise of AR technology brings a more intuitive and efficient way to provide warnings to onsite workers (Hou et al., 2020). AR technology visualises hazard information as holograms in the real world, which can help workers identify potential threats around them while viewing their surroundings. Kim et al. (2017) developed a proximity warning system for AR goggles, which provided a natural interaction to help workers identify hazards and avoid accidents. In this AR application, an arrow indicates the direction of the nearest construction vehicle, and texts show the safety level and the distance between the worker and the vehicle. This AR goggle-based system avoids any operational control. Onsite workers can identify the hazards around them with a quick glance. However, when evaluating the effectiveness of the system, the researchers placed the participants in the centre of a room instead of on a construction site.

Such simplification does not consider the holistic decision-making process of workers in complex construction scenarios. Wu et al. (2022) integrated Building Information Modelling (BIM) into an AR warning system that allows AR devices to provide warnings of multiple hazards, including explosions, falling from heights, and struck-by accidents. In their system, users can literally see the surrounding hazard areas with annotations in the real world. Their experiments only focused on the feasibility of their concept and paid no attention to the impact of the visual warnings on users. While the immersive AR experience provides users with hazard information in a natural and intelligent manner, it places a higher demand on the usability evaluation. First, because AR holograms are merged with the construction environment, workers can obtain AR prompts while scanning the environment. The act of obtaining AR prompts is natural and difficult to be captured. Second, the visual representation of AR prompts varies with the relative position between the worker and the hazard source. The effect of this variation on worker performance is difficult to quantify. Finally, the experiment involved a large amount of working memory which lasted for a very short period of time. It was difficult for workers to make precise recollections of the details after an experiment.

### 2.2 Eye-tracking in Construction Safety

Eye-tracking has already been widely used in the construction safety area for determining the hazard identification, situation awareness, and mental workload of construction workers. Pinheiro et al. (2016) verified the feasibility of eye-tracking technology in the construction safety area. They asked workers to view construction site images and analysed the collected Area of Interest (AoI) data. The results demonstrated that eye-tracking technology could clarify whether, when, and how a construction worker identifies hazards. Bhoir et al. (2015) used a head-mounted eye-tracking device to obtain fixation-related and saccade-related metrics of onsite workers and concluded that some workers did fail to notice hidden hazards or even hazard signs on construction sites. Hasanzadeh et al. (2018) explored the relationship between eye-tracking data and workers' situation awareness. The results demonstrated that eye-tracking could be used to quantitatively measure the extent to which workers are aware of hazards on a construction site. Dzeng et al. (2016) compared the search patterns of experienced and novice workers by analysing eye movement data and found that experienced workers showed fewer fixations, and their scan paths for assessing hazards were more consistent. Similarly, using eye-tracking data, Hasanzadeh et al. (2017b) demonstrated that hazard identification skills significantly affect the visual search strategies of construction workers. Hasanzadeh et al. (2017a) experimentally proved that there are significant differences in people's visual search strategies under different working memory loads. These studies focused on applying existing eye-tracking technology in construction safety and demonstrated that eye movements could reflect well on the performance of workers. The eye-tracking metrics and explanations in the existing studies are summarised in Table 1.

Table 1: Eye-tracking metrics and explanations in the reviewed articles.

| Category | Metric | Explanation |
|---|---|---|
| Fixation | Fixation duration | The time during which the eyes rest in a fixation |
| | Fixation count | The number of fixations that occur in an AoI. |
| | TFD | The total fixation duration in an AoI |
| Scan | Saccade speed | The speed at which the eyes move from an AoI to another |
| | Fixation sequence | The order in which the eyes move between AoI |

| | | |
|---|---|---|
| Physiological indicators | Pupil size | The average diameter of pupils within a certain period of time |
| | Blink frequency | The number of times an individual blinks eyes per minute |

Researchers have tried to introduce more technologies to increase the applicability of eye-tracking technology. Chew et al. (2018) introduced a method to model the gaze pattern of crane operators and indicated that varying visual support design elicits different gaze behaviour. They divided the operational scene of the crane into several AoI and recorded the duration and sequence of workers' gaze on these areas using eye-tracking devices. Then, a first-order Markov transition matrix was used to characterise the attention switch of each crane operator. Jeelani et al. (2018) integrated computer vision and eye-tracking into onsite workers' AoI in 3D spaces. They use a camera and a head-mounted eye-tracker to capture the workers' position and eye movement data. Then, an algorithm was proposed to combine these two data types for transforming workers' gaze positions from local coordinates to world coordinates in a 3D point cloud environment. This method can provide more objective data on gaze position and more accurately reflect whether a worker identifies hazards or not. However, each of these studies is based on a specific application scenario. The former requires the operator to be in a crane cabin, and the latter requires the construction site environment to be consistent with the point cloud reconstruction. These specific scenarios make it difficult to generalise and normalise the evaluation metrics for different systems.

### 2.3 Research Gaps and Research Objectives

Through the literature review, three limitations have been identified. First, as more modalities and natural interactions are introduced into warning systems, conventional metrics (i.e., warning accuracy and latency) are insufficient to evaluate the effectiveness of warning systems, especially when AR technology is introduced. We need to assess whether workers can perceive the existence of hazards, locate the hazards, correctly assess risks, and respond properly to the hazards (Wu et al., 2023). Second, there is a lack of standard quantitative methods to measure the impact of warning systems on onsite workers. The evaluation methods in existing studies were designed based on specific scenarios, leading to difficulties in comparing the efficiency and effectiveness of different systems. Although eye-tracking has the potential to objectively quantify the impact of warning systems on workers, related research is scant. Finally, existing studies have not applied AR warning systems in real construction situations to verify their usability. Experiments are less likely to reflect the actual impact of alert systems on onsite workers than case studies because the participants have a clear expectation that certain hazards will be present during the experiment. To address these limitations, we aim to develop an eye-tracking-based approach to measure the impact of warning systems on onsite workers. The objectives include:

1) Identifying metrics applicable for evaluating the performance of workers under the impact of safety warnings to reflect the process of perceiving, locating, and understanding hazards.
2) Developing a module that allows the impact of varying warning systems on workers in different scenarios to be automatically recorded, analysed, and visualised.
3) Verifying the feasibility of the proposed evaluation approach in a real construction environment.

## 3. Methodology

### 3.1 Metrics to Measure Hazard Identification Performance

The purpose of safety warnings is to help workers quickly perceive, locate, evaluate hazards, and respond to them appropriately. The metrics used to measure safety warning systems should reflect their performance in these aspects. Accordingly, we categorised the information that the warning system can provide into three categories: the existence of hazards, the locations of hazards, and other details. Indicating the existence of a hazard to workers is the basic function of a safety warning system. To measure the effectiveness of the hazard existence information, we follow the method used in existing studies, which considers a warning valid if a worker perceives the existence of a hazard within five seconds after a hazard warning is presented. The ratio of valid warnings to the total warning number is set as Valid Warning Rate (VWR). The duration of receiving information (DRI) represents the time spent being noticed and receiving information delivered by warning systems. Workers' visual attention significantly changes when they notice a hazard. Such changes can take the form of viewing warnings or scanning the environment for hazards, and eye-trackers can capture them to determine VWR and DRI.

Since warning information is stored in workers' working memory, we introduced a working memory theory to this study (Peterson and Peterson, 1959). In this theory, the amount of information that can be retained in working memory gradually decreases with time. The reduction starts quickly and gradually slows down. Generally, the retained information drops below 15% after 18 seconds, indicating that workers may need to recheck the warning. Therefore, two metrics were used to measure the ability of the warning system to help workers locate hazards: the number of rechecks (Recheck Count, RC) and the Duration of Locating Hazard (DLH). To measure the extent to which the warning system helps workers assess hazards for decision-making, we set the Duration of Decision-Making (DDM) as a metric. This refers to the time spent observing hazards. Additionally, the total time was also set as a metric.

### 3.2 Measurement Module Development

In order to obtain the value of these metrics, the following functions are required: 1) tracking the location of hazards, 2) tracking the location of workers, 3) tracking the 3D AoI of onsite workers, 4) recording the time and duration at which workers check warnings (including viewing AR content), and 5) automatically generating metric values for each warning. Two modules were developed to implement the above functions, namely, an object detection module and an eye-tracking module. The object detection module uses object detection and coordinate transformation algorithms to track onsite workers and dynamic hazard areas. It can also invoke the positioning Application Programming Interface (API) of AR devices to determine the users' position. The eye-tracking module interfaces with the API of eye-tracking devices to continuously monitor the direction of the worker's gaze and calculate the point where the line of sight makes contact with AR holograms or real objects. When a worker views an AR warning or looks at a danger zone, the system automatically triggers the appropriate function to record coordinate data for generating the metric values, scanpaths, and heat maps. These two modules can be integrated into applications developed by the Unity game engine.

### 3.3 Case Study

To test the impact measurement method, we deployed an AR warning system containing the developed modules on a Microsoft HoloLens 2, and had a construction cadet wear it to perform a three-hour construction activity at a construction site. The activities involved moving materials, pipe laying, formwork, rebar tying, and other works. Workers were required to perform these tasks while ensuring their own safety. The AR goggles deployed two types of warnings, namely, AR visual warnings developed by Wu et al. (2022) and beeping alarms proposed by Teizer (2015). The visual warning provided a holographic arrow pointing to hazard sources and a holographic warning sign displaying the hazard types, as shown in Figure 1. These two kinds of warnings were switched every 15 minutes to test the performance of the worker under different warnings. To ensure safety, we eliminated possible hazards at the construction site and used several methods to simulate the unsafe environment. For example, when workers moved oxygen cylinders, the areas near acetylene cylinders became hazardous areas. So empty bottles labelled acetylene were set instead of real acetylene bottles. Similarly, a driver's blind spot can lead to a struck-by accident, so drivers on the construction site were only allowed to get in vehicles to represent the vehicles were about to move rather than actually drive them.

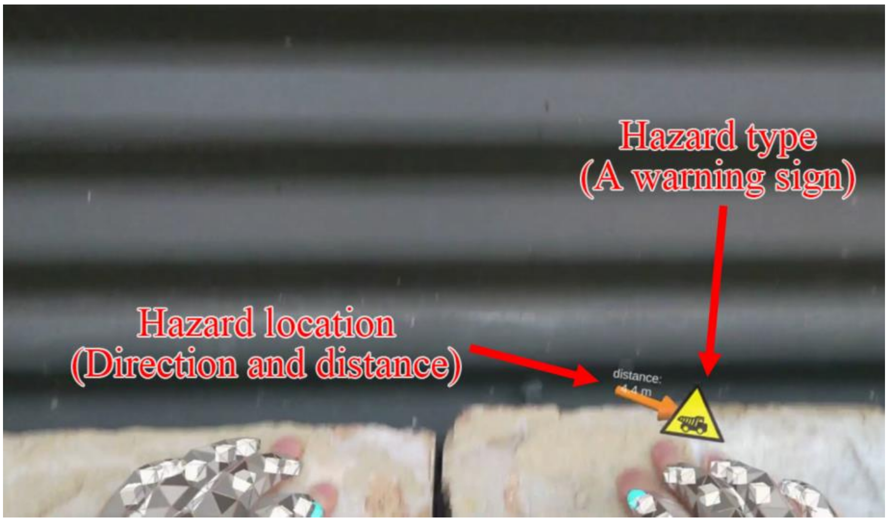


Figure 1: An example of the AR visual warnings

## 4. Results and Discussion

The case study demonstrates that the developed module effectively captures the eye movement of onsite workers, with an average frame rate of 59.99Hz and a maximum recording interval of 26.2ms. During the study, the participant received a total of seven warnings, consisting of four visual warnings and three beeping alarms. The VWRs for both types of warning were 100%, and the participant's performance under the impact of safety warnings is shown in Table 2.

Table 2: Eye-tracking results of the case study.

| Warning type | Warning No. | Hazard Type | RC | DRI (ms) | DLH (ms) | DDM (ms) | Total time (ms) |
|---|---|---|---|---|---|---|---|
| Visual | 1 | Struck-by | 0 | 1001.13 | 749.67 | 2232.66 | 3983.46 |
| | 2 | Struck-by | 0 | 836.72 | 826.45 | 2563.59 | 4226.76 |
| | 3 | Exploration | 1 | 1593.27 | 1426.72 | 3147.46 | 6167.45 |
| | 4 | Electrocution | 0 | 982.03 | 743.14 | 1841.33 | 3566.50 |
| | Ave. | - | 0.25 | 1103.29 | 936.50 | 2446.26 | 4486.04 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Auditory | 5 | Struck-by | - | 482.02 | 3265.19 | 3446.10 | 7193.31 |
| | 6 | Exploration | - | 592.08 | 1515.23 | 3857.13 | 5964.44 |
| | 7 | Struck-by | - | 518.45 | 2729.01 | 2385.92 | 5633.38 |
| | Ave. | - | - | 530.85 | 2503.14 | 3229.72 | 6263.71 |

Time-related metrics show that AR visual warnings have better overall effectiveness than beeping alarms. The participant took about twice as long to identify visual warnings (1103.29ms on average) compared to detecting beeping alarms (530.85ms on average). However, viewing visual information significantly reduced the time spent locating hazards and making decisions. After checking the visual warnings, the participant only needed an average of 936.50ms to locate hazard sources, which is a significant improvement compared to the beeping alarm (2503.14ms). The heat maps and scanpaths reveal that the participant needed to scan the environment to identify hazards due to the lack of exact location information. In the worst case, the hazard source was located at the fifth fixation, as shown in Figure 2. In contrast, the participant could quickly locate the hazard source in a similar situation after viewing the AR holograms. Additionally, the participant spent less time making decisions after receiving AR visual warnings (2446.25ms for AR visual warnings vs. 3229.72ms for auditory warnings). Overall, using AR visual warnings saved 28.38% of the time compared to using beeping alarms.

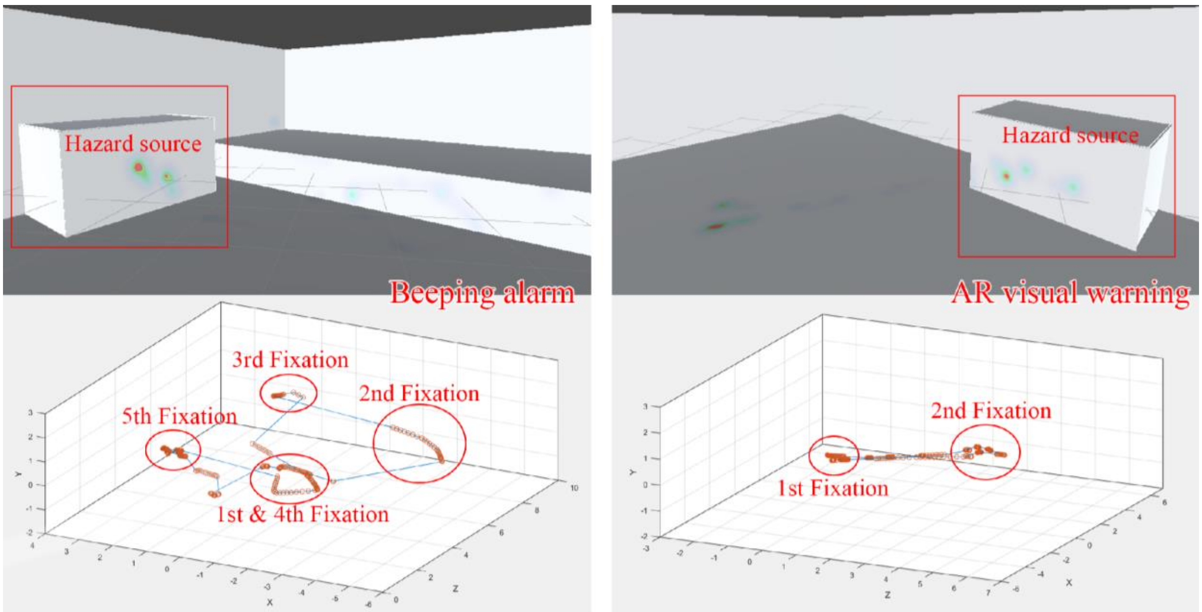


Figure 2: Heat maps and scanpaths generated from eye-tracking results.

It is worth noting that a participant viewed an AR hologram twice under an explosion hazard situation. The participant did not realise that he had such behaviour until he viewed his eye-tracking data. This is probably because the amount of information displayed by the AR visual warnings in this case exceeded the capacity of the participant's working memory. The participant subconsciously rechecked the warning message and extracted key information to consolidate his memory. Therefore, visual warnings should be designed with the appropriate information to avoid working memory overload.

## 5. Conclusion

This study proposes a method to quantify the impact of warning systems on onsite workers through the innovative introduction of eye-tracking technology. The proposed method has been proven effective in reflecting workers' performance under the influence of safety warnings. By employing eye-tracking technology, the fixation positions, fixation durations, and scanpaths of workers can be recorded automatically. These eye movement data can be summarised as six quantitative metrics, namely, VWR, RC, DRI, DLH, DDM, and total time, to measure the impact of a safety warning on onsite workers. The proposed approach is compatible with eye-tracking devices and AR goggles, making it applicable to various warning types, including auditory, vibration, and AR visual warnings. This enables objective and quantitative comparisons between different systems. In addition, the method has been proven to be applicable to real construction sites so that researchers can use it to measure and analyse the impact of warning systems in real-life use cases.

## References


Abdelhamid, T. S. & Everett, J. G. (2000). Identifying root causes of construction accidents. *Journal of construction engineering and management,* 126**,** 52-60.

Bhoir, S. A., Hasanzadeh, S., Esmaeili, B., Dodd, M. D. & Fardhosseini, M. S. (2015). Measuring construction workers attention using eye-tracking technology.

Chew, J. Y., Ohtomi, K. & Suzuki, H. (2018). Glance behavior as design indices of in-vehicle visual support system: A study using crane simulators. *Applied ergonomics,* 73**,** 183-193.

Dzeng, R.-J., Lin, C.-T. & Fang, Y.-C. (2016). Using eye-tracker to compare search patterns between experienced and novice workers for site hazard identification. *Safety science,* 82**,** 56-67.

Hareide, O. S. & Ostnes, R. (2016). Comparative study of the Skjold-class bridge-and simulator navigation training.

Hasanzadeh, S., Dao, B., Esmaeili, B. & Dodd, M. D. (2017a). Measuring the impact of working memory load on the safety performance of construction workers. *Computing in civil engineering 2017.*

Hasanzadeh, S., Esmaeili, B. & Dodd, M. D. (2017b). Impact of construction workers' hazard identification skills on their visual attention. *Journal of construction engineering and management,* 143**,** 04017070.

Hasanzadeh, S., Esmaeili, B. & Dodd, M. D. (2018). Examining the relationship between construction workers' visual attention and situation awareness under fall and tripping hazard conditions: Using mobile eye tracking. *Journal of construction engineering and management,* 144**,** 04018060.

Hou, L., Wu, S., Zhang, G., Tan, Y. & Wang, X. (2020). Literature review of digital twins applications in construction workforce safety. *Applied Sciences,* 11**,** 339.

Huang, Y., Hammad, A. & Zhu, Z. (2021). Providing proximity alerts to workers on construction sites using Bluetooth Low Energy RTLS. *Automation in Construction,* 132**,** 103928.

Jeelani, I., Han, K. & Albert, A. (2018). Automating and scaling personalized safety training using eye-tracking data. *Automation in Construction,* 93**,** 63-77.

Kanan, R., Elhassan, O. & Bensalem, R. (2018). An IoT-based autonomous system for workers' safety in construction sites with real-time alarming, monitoring, and positioning strategies. *Automation in Construction,* 88**,** 73-86.

Kim, K., Kim, H. & Kim, H. (2017). Image-based construction hazard avoidance system using augmented reality in wearable device. *Automation in Construction,* 83**,** 390-403.

Li, W.-C., Zhang, J., Le Minh, T., Cao, J. & Wang, L. (2019). Visual scan patterns reflect to human-computer interactions on processing different types of messages in the flight deck. *International Journal of Industrial Ergonomics,* 72**,** 54-60.

Martinez-Marquez, D., Pingali, S., Panuwatwanich, K., Stewart, R. A. & Mohamed, S. (2021). Application of eye tracking technology in aviation, maritime, and construction industries: a systematic review. *Sensors,* 21**,** 4289.

Peterson, L. & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of experimental psychology,* 58**,** 193.

Pinheiro, R., Pradhananga, N., Jianu, R. & Orabi, W. (Year) Published. Eye-tracking technology for construction safety: A feasibility study. ISARC 2016-33rd International Symposium on Automation and Robotics in Construction, 2016.

Sakhakarmi, S., Park, J. & Singh, A. (2021). Tactile-based wearable system for improved hazard perception of worker and equipment collision. *Automation in Construction,* 125**,** 103613.

Teizer, J. (2015). Wearable, wireless identification sensing platform: self-monitoring alert and reporting technology for hazard avoidance and training (SmartHat). *Journal of Information Technology in Construction (ITcon),* 20**,** 295-312.

Teizer, J., Allread, B. S., Fullerton, C. E. & Hinze, J. (2010). Autonomous pro-active real-time construction worker and equipment operator proximity safety alert system. *Automation in Construction,* 19**,** 630-640.

Wu, S., Hou, L., Chen, H., Zhang, G., Zou, Y. & Tushar, Q. (2023). Cognitive ergonomics-based Augmented Reality application for construction performance. *Automation in Construction,* 149**,** 104802.

Wu, S., Hou, L. & Zhang, G. (Year) Published. Integrated application of BIM and eXtended reality technology: a review, classification and outlook. Proceedings of the 18th International Conference on Computing in Civil and Building Engineering: ICCCBE 2020, 2021. Springer, 1227-1236.

Wu, S., Hou, L., Zhang, G. & Chen, H. (2022). Real-time mixed reality-based visual warning for construction workforce safety. *Automation in Construction,* 139**,** 104252.