Impact of Mobile Augmented Reality System on Cognitive Behavior and Performance During Rebar Inspection Tasks

Ali Abbas*, JoonOh Seo, and MinKoo Kim*****

*Ph.D. Student, Department of Building and Real Estate, The Hong Kong Polytechnic University, 11 Yuk Choi Rd., Hung Hom, Kowloon, Hong Kong SAR. Email: ali.abbas@connect.polyu.hk.

**Corresponding Author; Assistant Professor, Department of Building and Real Estate, The Hong Kong Polytechnic University, 11 Yuk Choi Rd., Hung Hom, Kowloon, Hong Kong SAR. Email: joonoh.seo@polyu.edu.hk.

***Assistant Professor, Department of Building and Real Estate, Hong Kong Polytechnic University, 11 Yuk Choi Rd., Hung Hom, Kowloon, Hong Kong SAR. E-mail address: minkoo.kim@polyu.edu.hk.

Abstract

Mobile augmented reality (MAR) enhances the real world through the superimposition of computer-generated information while not interfering with their users' mobility, having great potential to support various construction tasks. However, such information may lead to cognitive overload and thus could lead to adverse effects on the performance of tasks. Also, the narrowing of a user's field of view that comes with MAR use could limit his/her ability to notice events in their surroundings. Therefore, it is important to understand how MAR use affects cognitive behavior, as well as task and safety performance for better design and applications of MAR in construction. As a preliminary investigation, this study conducted laboratory simulations of rebar-inspection tasks, and compared the cognitive load (CL), task performance (TP) and situational awareness (SA) of users of two types of MAR system – i.e., head-mounted and handheld – against those of inspectors using traditional paper-based methods. In particular, participants' CL was measured with the NASA-TLX; their TP, by completion time and the number of rebars correctly detected; and their SA, with Taylor's SART. Based on the results,

we discuss the impact of the MAR system on rebar-inspection tasks from both cognitive and safety perspectives.

Keywords: Mobile augmented reality, cognitive behavior and performance, cognitive load, situational awareness, rebar inspection

INTRODUCTION

Augmented reality (AR) is a technology for enhancing the real world by superimposing computer-generated information such as computer graphics, text, or sound onto real-world scenes (Kalawsky et al., 2000). Architecture, engineering and construction (AEC) industry stakeholders are embracing its potential applications at various project stages, including visualization during the design stage (Alsafouri and Ayer, 2019); safety management/inspection during construction (Heinzel et al., 2017; Olsen et al., 2019); and information access (Irizarry et al., 2013) and evaluation for maintenance (Ammari and Hammad, 2014) during the facility-management stage (Rankohi and Waugh, 2013). One prominent benefit of using AR at construction sites is that it enables construction stakeholders to review construction drawings at full, i.e., 1:1 scale, and thus identify errors that might not otherwise be spotted (Agarwal, 2016). For example, installation of a structural steel column requires not only the placement of its base in a specific location, but also a critical 3D assessment of its vertical alignment. Thus, AR can help prevent steel-column installation errors and save inspection time, since each object in its superimposed model is uniquely referenced to a unified system of coordinates, eliminating the possibility of errors accumulating across different sets of reference materials drawn at multiple scales (Dunston, 2009). Also, by marrying spatial data to real-world physical objects and locations, AR supports construction tasks such as a layout task, the process whereby relevant points in a construction space are earmarked for future work, by strongly leveraging its users' spatial cognition and memory (Chalhoub et al., 2019). As such, AR assistance for cognitive-based construction tasks such as assembly work (Lei et al., 2013), point layout (Chalhoub et al., 2019) and inspection (Zhou et al., 2017) could reasonably be expected to reduce both mental workload and task-completion time.

While various types of AR devices and systems have been developed, mobile AR (MAR) systems are increasingly prominent, as they allow AR to be moved from the laboratory onto actual construction sites (Izkara et al., 2007). MAR can be divided into two main categories – handheld devices such as tablets, and wearable devices like smart glasses and head-mounted displays (HMDs) – both of which afford their users high mobility and anytime/anywhere management of spatially registered information. Some previous AEC-focused research on AR has looked at how to apply it to and through mobile devices, such as for registration of virtual objects, real-time tracking, and calibration (Bae et al., 2013; Kopsida and Brilakis, 2016; Kwon et al., 2014). Unsurprisingly perhaps, the usefulness and technical advancement of MAR have taken center stage in such research, which in most cases has ignored that the AR environment could create perceptual issues, including but not limited to field-of-view, registration, and depth-perception errors (Dey et al., 2018). These issues, in turn, could severely affect users' cognition, performance, and comprehension of augmented content (Kruijff et al., 2010). In addition, the reference frame of AR information is critical to the cognitive functioning needed to understand one's surroundings when using MAR (Li and Duh, 2013). Nevertheless, previous studies' proposed MAR designs have not given due consideration to these issues, and no specific MAR design guidelines exist (Li and Duh, 2013). These absences necessitated the current investigation of how cognitive factors and corresponding task and safety performance could be affected by MAR environments.

In this regard, we aim to understand the effects of two distinct types of MAR (i.e., handheld and head-mounted systems) on construction professionals' cognitive load (CL), task performance (TP), and situational awareness (SA), relative both to each other and to paper-based techniques. To achieve this research objective, we conducted experimental studies of a rebar-inspection task that is not only information-intensive, but also cognitively demanding, at construction sites. Specifically, our three participant groups were given the task of inspecting rebar for a concrete slab using MAR on a tablet, MAR on Microsoft HoloLens, and traditional drawings. TP was measured using task-completion time and error-identification rate; CL was measured using the National Aeronautics and Space Administration's Task Load Index (NASA-TLX); (Hart, 2006) in a laboratory setting; and SA was measured using the Situation Awareness Rating Technique (SART); (Taylor, 1990) and by simulating a construction site-like environment in a laboratory.Based on the result, we discussed participants' TP , CL, and SA of the surrounding environment in traditional drawings and MAR -assisted rebar inspection.

LITERATURE REVIEW

Application of Mobile Augmented Reality in Construction

MAR's known and potential capabilities are attracting AEC industry stakeholders to embrace its use during various stages of their projects. Wang (2007) used ARTag tracking markers and ARToolKit software to plan construction worksites through AR, and highlighted that traditional 2D paper media were less effective than MAR when it came to understanding both spatial constraints and resource-allocation strategies. Woodward and Hakkarainen (2011) proposed a MAR for construction-site visualization and interaction with complex 4D building-information models, and Kim et al. (2013), construction job-site defects monitoring using MAR and computer vision-based algorithms. Kwon et al. (2014) used a MAR with ARToolkit to automatically detect dimensional errors and omissions on the worksite and found it easier to use for this purpose than the manual-based defect management process. Kopsida and Brilakis (2016) used a makerless building information modeling (BIM) registration method for MAR-based inspection and reported that it reduced inspection time by providing the inspector with instantaneous access to the information stored in the BIM. Zaher et al. (2018) developed two MAR applications that allow their users to update the progress of construction-site activities, which can be used through implementing a 4D 'as-planned' phased model integrated with an augmented video showing real or planned progress. Alsafouri and Ayer (2019) investigated the feasibility of wearable and handheld MAR systems for industry practitioners in design and constructability-review sessions, and found that both allowed their users to 'walk through' and interact with virtual environments, facilitating their decision-making, problem-solving, and creation of design alternatives. Olsen et al. (2019) used MAR through wearable Microsoft HoloLens device for inspecting missing or misaligned embeds, sleeves and penetrations in concrete and masonry construction, and found that HoloLens speeded up the locating of the embeds, which are hard to represent in 2D drawings. Lastly, Lamsal and Kunichika (2019) developed an AR system specifically for adaption to MAR via iPads and other tablet computers using Vuforia and AR markers, and tested it on the rebar construction phase of a 13-story steel building in Japan, reporting its strong potential to increase productivity.

As the above discussion suggests, the MAR systems developed to date have been very diverse, with features including touchscreens providing virtual keyboards and onscreen buttons, integrated cameras, wireless connectivity, global positioning system capabilities, and computer-generated data displays (Alsafouri and Ayer, 2019). Yet, while all the studies cited above have endorsed the use of MAR applications for at least one construction task, little research has focused on MAR's impacts on AEC-industry users' perception and/or cognitive behavior, or how such impacts may be linked to AEC task performance. The present study is intended to fill those research gaps.

Potential Impact of Mobile Augmented Reality on Cognitive Behavior and Corresponding Task and Safety Performance

Human cognitive-behavioral research focuses on understanding how mind, brain, and body interact, through observation of human cognitive behavior such as CL, SA, perceptual processing, and information processing (Curtin and Ayaz, 2017). While a number of definitions of CL exist, this study adopts Brunken et al.'s (2003) view that it comprises the amount of mental effort one expends during information processing. According to Doswell and Skinner's (2014) CL theory, human working memory can only simultaneously handle an average of seven (plus or minus two) disconnected items; and thus, cognitive overload tends to occur when human working memory is forced to process larger amounts of information quickly. As such, the amount of information that needs to be handled can significantly affect a person's task performance. From a cognitive perspective, all the major tasks in the AEC industry involve information-intensive processes, so under such conditions, MAR interfaces could overload the user with information, such that important cues from their actual environments could be missed. Previous studies on MAR systems to support surgical procedures (Doswell and Skinner, 2014) and procedural tasks (Baumeister et al., 2017) have reported that the use of MAR systems in the complex environment could lead to increase cognitive burden. In addition, Li and Duh (2013)'s study raised the cognitive issues based on the findings of existing literature and explained that an excessive amount of information, its representation, placement, and view combination visualization techniques of MAR assisted system such as zooming and panning to understand the meaning of detailed information could impact the user's cognitive functioning. Considering that the nature of construction sites is dynamic and complex, it is more expected that an excessive amount of information and its placement in the MAR assisted system could increase the visual-processing and information interpretation issues and have negative impacts on construction workers CL, TP, and SA.

In general terms, SA consists of being aware of what is happening around you. More specifically, Endsley (1988a, p. 97) defined it as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future." MAR has a strong inherent potential to enhance visual perception via superimposition of information and thus has been argued to enhance overall SA (Lukosch et al., 2015). However, in most cases, AR environments have been found to cause perceptual issues during the visual processing and interpretation of information, affecting field of view, registration, depth perception, and so on, which in turn negatively impact the user's cognition, performance, and comprehension of augmented content (Kruijff et al., 2010). In addition, various studies (Lindblom and Thorvald, 2014; Lyell et al., 2018; Paas et al., 2004) have looked at the general relationship between cognitive issues and performance (see Fig. 1). Specifically, Paas et al. (2004) and Lindblom and Thorvald (2014) found that too little CL (underload) as well as too much CL (overload) could lead to performance issues. For example, cognitive underload can occur when a user heavily relies so heavily on a system during tasks that he/she may lose interest in them, leading to more task-related errors. But at the other extreme, the amount of information coming from a system can surpass and overwhelm human processing capacity. These insights led Mendel and Pak (2009) to argue that user performance could be increased by reducing CL during information-intensive tasks.

In conclusion, prior studies indicate that the amount of information provided to AR system users can influence their performance, and should be carefully considered, with too much and too little information both being problematic. Given that any MAR system can only achieve optimal performance when it provides an appropriate amount of information, it is critically important to gauge users' CL in specific MAR environments, as well as how variations in that CL relate to their performance.

Because a construction site is a complex, dynamic environment, failure to address AR users' visual-processing and information-interpretation issues could have serious negative effects on worksite safety (Bhandari et al., 2018). However, previous proposals for MAR systems for use in AEC have not fully considered these issues, and there are no specific design guidelines that take account of how MAR environments may affect CL, TP and SA (Li and Duh, 2013). In-depth understanding of how MAR can affect its users' cognitive behavior and performance would be helpful in the creation of such guidelines, and therefore to the design of safer and more effective MAR systems for AEC use.

Cognitive Load, Task Performance, and Situational Awareness Measures

CL is commonly measured using one or more of four broad sets of techniques: subjective, performance, physiological and behavioral (Khawaja et al., 2014). The subjective techniques primarily include gathering data directly from subjects, who rate their own CL on a Likert-type scale. The most reliable subjective CL results have generally been attained using the NASA-TLX (Hart, 2006). The performance-based CL measurement technique, on the other hand, assesses subjects' performance while a task is being carried out: for example, using task-completion time, critical errors, false starts, speed and/or correctness (Paas et al., 2003). The physiological approach, meanwhile, relies on changes in human cognitive functions being reflected physiologically, e.g., through brain activity, eye movement, or heart rate (Joseph, 2013). And lastly, behavioral measures can provide nonintrusive, objective, and implicit analyses of individuals' CL, as they are based on data collected during task completion without the participants' prior knowledge. Some commonly used behavioral measures of CL include speech features (e.g., pitch, prosody) and linguistic features (e.g., pauses, patterns of language) (Khawaja et al., 2014). Across all CL measurement techniques, however, the NASA-TLX is one of the easiest to use, least expensive, most reliable, and most sensitive to small variations in workload (Bhandary et al., 2016; Dadi et al., 2014; Hou et al., 2013).

Cognitive TP measures are based on the assumption that the mental workload of an individual interacting with a particular system or interface during the performance of a particular task is a good indicator of CL (Lee et al., 2018). Examples of cognitive TP metrics include reaction time to a secondary task, task-completion time, and error rate (Longo, 2018).

In their reviews of SA measurement techniques, (Salmon et al., 2006; Salmon et al., 2009) categorized past approaches into five general types, including (1) physiological methods such as eye-tracking and electroencephalograms (EEGs); (2) performance-based methods for example mission success or failure, hazard detection etc; (3) self-rating methods, such as SART

(Taylor, 1990), the Crew Awareness Rating Scale (McGuinness and Foy, 2000) or the Mission Awareness Rating Technique (Matthews and Beal, 2002) (4) observer-based rating methods like the Situation Awareness Behavioral Rating Scale (Matthews et al., 2005); and (5) freeze-probe methods, such as SA global-assessment techniques (Endsley, 1988b). Although, there are both advantages and disadvantages of each technique, among these, SART is widely acknowledged as non-intrusive, inexpensive, easy to perform, and simple to analyze (Endsley and Garland, 2000; Endsley et al., 1998; Stanton et al., 2005). Its three key dimensions – i.e., understanding of the situation, demands on attentional resources, and supply of attentional resources – together provide comprehensive measurement of individuals' SA (Hasanzadeh et al., 2018; Naderpour et al., 2016; Salmon et al., 2009).

RESEARCH METHODOLOGY

To achieve this study's research objective, as shown in Fig. 2, the participants we recruited were assigned to one of three rebar-inspection groups: one using paper-based inspection methods, another using tablet-based MAR, and a third, HoloLens-based MAR. Two experiments were performed. Experiment I assessed the respective impacts of the traditional paper medium and each type of MAR on the participants' TP (as measured by completion time and number of errors) and CL as measured by using NASA-TLX (Hart, 2006) in a laboratory environment. Then, Experiment II added a simulated construction site to the laboratory environment and assessed how each of the three inspection modalities affected individuals' awareness of the surrounding environment during inspection tasks, and their overall impact on CL, TP, and SA. The task again consisted of rebar inspection, albeit with a different slab rebar framework to minimize learning effects; and each participant used the same inspection modality that he/she had used in Experiment I. To assess the impact of construction-safety conditions on Experiment II's results, we performed inter-group comparisons of CL and TP, and also measured each participant's SA using SART (Taylor, 1990). The procedures of both experiments are explained below in greater detail.

Participants

A sample of 45 Ph.D. students from the Department of Building and Real Estate at the Hong Kong Polytechnic University was recruited for the two experiments. All participants had previously taken multiple classes related to construction project management and had some professional construction-industry experience, and thus were familiar with rebar inspection. They were randomly divided into three groups of 15, each of which would perform its rebar inspections using the same modality (i.e., paper, tablet MAR, or HoloLens MAR) across both experiments. Before experimental sessions, we provided clear and concise instructions to each participant by using organized materials regarding the experimental procedures and provided multiple training sessions on how to use MAR devices for rebar inspection. In addition, to avoid the potential response bias during the post-experiment surveys, we made questions concise and easy to understand, and informed participants that survey data would be strictly used for research purposes only on an anonymous basis.

Task Overview

In both experiments, all participants played the role of a construction inspector tasked with checking for the following eight types of reinforcement errors: (1) spacing between rebars, (2) missing rebars, (3) extra rebars, (4) insufficient rebar cover at the side face, (5) insufficient rebar cover at the bottom face, (6) incorrect number of anchorage bars, (7) insufficient length of anchorage bars, and (8) bars incorrectly tied and supported. In all, 20 errors were intentionally placed in the rebar framework to be inspected, as shown in Fig. 3.

Experimental Procedure

During the paper-based inspection session of each experiment, the participants were asked to find rebar errors of each of the eight types given above by comparing the physical rebar framework against a drawing, as shown in Fig. 4 (left), using a tape measure if they wished. The second group of participants performed the same task using a tablet that, when pointed at the physical rebar framework, showed a 3D rebar model superimposed on it, as shown in Fig. 4 (middle). This 3D rebar model had first been drawn in SketchUp and then integrated with SketchUp Viewer, a tablet AR app (https://www.sketchup.com/products/sketchup-viewer). The third group of participants performed the same task while wearing Microsoft HoloLens headsets that showed a 3D rebar model superimposed on physical rebar framework, as shown in Fig. 4 (right). This second 3D rebar model was also first drawn in SketchUp, but then integrated with Trimble Connect, a HoloLens-specific AR app (https://mixedreality.trimble.com/). Participants in all three groups were instructed to perform the inspection task as fast and accurately as possible, with their respective inspection speeds and numbers of errors both being collected in real-time. NASA-TLX was then used at the end of each experiment to measure their CL.

One week after Experiment I, we conducted a very similar experiment, with the same groups using the same inspection modalities, but a different rebar model, and with a more realistic simulation of a construction environment within the laboratory. Specifically, this environment was designed to expose the participants to realistic construction scenarios as a test of their SA: with recorded sounds of construction equipment played at accurate volumes, and a person employed to drive a laden forklift trolley near each participant during his/her inspection task. During this experiment, the same techniques as in Experiment I were used to measure participants' TP and CL, while SART was used at the end of the experiment to measure their SA.

Measurements

To measure cognitive load, we used NASA- TLX method that has been widely used for measuring cognitive load (Bhandary et al., 2016; Dadi et al., 2014; Hou et al., 2013). The original NASA- TLX contains six items (mental demand, physical demand, performance, temporal demand, effort, and frustration level). However, physical demand – defined as how much physical activity is required during a task – was not deemed relevant to our research, and so was omitted from the version of NASA-TLX that was used. One of the remaining five items, performance, could have been measured directly; however, as used in the NASA-TLX, it incorporates non-objective factors such as level of satisfaction, self-esteem, and motivation, and we retained it for that reason. Therefore, based on mental demand, performance, temporal demand, effort, and frustration level, participants in each experiment were rated on a scale from 1=Low to 5=High, as shown in Table 1.

While TP was measured objectively, as a combination of (1) the actual amount of time a participant took to complete his/her assigned inspection task in a given experimental session, and (2) the number of rebar errors that he/she correctly identified during that session.

Finally, to measure SA we used the SART method. It is a well-known post-trial subjective rating technique for the assessment of a participant's SA, further details of which are shown in Table 2. SART was completed by our participants at the end of Experiment II using a five-point Likert scale ranging from 1=Low to 5=High. The original SART instrument contains 10 items covering the environment's (1) information quantity, (2) information quality, and (3) the participant's familiarity with it ; (4) the instability (5) the variability of the prevailing situation and (6) complexity.; (7) arousal, (8) concentration, (9) division of attention, and (10) spare mental capacity. However, these 10 items can be grouped into three major dimensions: i.e., understanding of the surrounding situation (U), demand on attentional resources (D), and supply of attentional resources on the surrounding situation (S), where U is the sum of items (1), (2), and (3); D, is the summation of items (4), (5) and (6); and S, is the summation of items (7) through (10). A person's overall SART score can then be calculated as SA=U-[D-S].

RESULTS

Before analyzing the data in detail, we first performed a Shapiro-Wilk test, a widely used method of testing data normality in sample sizes smaller than 50 (Ahad et al., 2011; Mishra et al., 2019). The common alpha value for testing normality (i.e., 0.05) was used in conducting this test, and if the *p*-value produced by the test is lower than the accepted value, then we can conclude that the data are not normally distributed (Darko and Chan, 2018). All the *p* values produced by the Shapiro-Wilk testing of the present study's data were 0.00, indicating that such data were not normally distributed. Therefore, non-parametric tests – which are considered suitable for non-normally distributed data – were used for the remainder of our analyses. Non-parametric Kruskal-Wallis H can be used to assess statistically significant differences among three or more independently sampled groups (McKight and Najab, 2010), and therefore was chosen for use with both the Experiment I and Experiment II data to identify any statistically significant differences among the paper, tablet and HoloLens users.

Experiment I: General Comparison among Inspection Modalities

As the purpose of Experiment I was to assess how traditional paper-based inspection and the two focal types of MAR would affect the participants' CL, the Kruskal-Wallis H test was conducted first, as shown in Fig. 5. Its results indicated that paper-based inspection was the most cognitively demanding of the three modalities, and HoloLens the least, though differences among them were not statistically significant. Then, a detailed comparison was made of the three inspection groups' NASA-TLX data. As Fig. 5, indicates, users of both MAR systems perceived lower CL than the participants using the paper-based inspection method did. Again, however, the mean differences were found to be non-significant ($p>0.05$).

Next, the Kruskal-Wallis H test was applied to the Experiment I data on users' average completion times (Fig. 6) and error-identification rates (Table 3). As shown in Fig. 6 which presents a comparison of completion times across the three experimental groups, the paper-based group, at 11.85 minutes, took significantly longer than either of the two MAR-assisted groups (p<0.05). However, there was no statistically significant difference between the completion times of the tablet-based and HoloLens-based MAR groups (6.17 and 6.59 minutes, respectively; p>0.05).

Each group's error-identification rate was analyzed through the Kruskal-Wallis H test, as shown in Table 3. There were no statistically significant differences among the three groups' mean performance at identifying missing-bar and extra-bar errors. However, statistically significant differences did emerge between both MAR groups, on the one hand, and the paper-based group, on the other, when it came to identifying spacing, side-cover, bottom-cover, bar-number, length, and tying/support errors (p<0.05), with the paper-based group performing significantly better in these areas. And overall, out of 20 errors that were intentionally placed in the physical rebar framework, an average of 13.5 were correctly identified by the paper-based group, as against 9.5 by HoloLens users and just 9.1 by tablet users; and this difference was also found to be statistically significant (p<0.05).

Experiment II: Relationships between Safety Conditions and Inspection Modalities

The Kruskal-Wallis H test was performed on the Experiment II data to see how the addition of realistic construction sounds and potentially dangerous environment affected the participants' CL, TP, and SA. Although no significant mean difference in CL was found across the two experiments (as shown in Fig. 7), average CL for all three inspection groups was higher in Experiment II than in Experiment I.

As shown in Fig. 8, we also found that average completion time for each inspection modality was higher in Experiment II than in Experiment I. However, this mean difference was found to be statistically significant only for the HoloLens group. As indicated in Table 4, the

Experiment II data also showed that fewer errors were identified by the tablet and HoloLens users than by the traditional-inspection group. The latter group was also exceptional in that the increased environmental noise and hazard levels had no marked negative impact on its error-identification performance. However, no statistically significant overall difference in error identification was found between Experiment I and Experiment II.

Finally, we examined the inspection-group SART scores from Experiment II through the Kruskal-Wallis H test. Table 5 presents the cumulative mean SART values, along with their SDs, Kruskal-Wallis H values, and significance levels (*p*). For this purpose, we first grouped the 10 SART items into the three main dimensions U, D, and S, as described above. There were significant mean differences in two of these three SART dimensions, i.e., D and S ($p<0.05$). While no such significant difference was found for the third dimension, U, the paper-based inspection modality still had a higher U (9.8) than either its tablet-based (9.53) or HoloLens-based counterpart (8.86). The cumulative average values of D were also found to be highest in the paper-inspection group (10.31, vs. 9.26 for the tablet group and 9.18 for the HoloLens group). Lastly, the cumulative average values of S were highest for the paper-inspection group (13.39). Total SART score, calculated using the formula Situational Awareness=Understanding-[Demand-Supply], was higher on average in the paper-based inspection modality (12.88) than in either the tablet (11.53) or HoloLens modality (10.93); however, these differences were not statistically significant ($p>0.05$).

DISCUSSION

This study compared the impact of two popular types of MAR (i.e., handheld and head-mounted systems) on CL, TP and SA. Through Experiment I, we revealed that the rebar-drawing information provided by superimposed computer imagery in both MAR systems helped to decrease their users' CL, as compared with traditional paper-based inspection. Also, we found that the paper-based group took more time to complete their inspection task than either of the MAR-assisted groups. However, because of perception issues associated with both MAR systems, notably involving depth and registration, the paper-based group identified more errors than either of its MAR-assisted counterparts. Then, Experiment II established that a more realistic construction-site environment increased the cognitive demand on the subjects and lowered their TP; and that the same environment also negatively impacted SA across all three dimensions of the SART.

In terms of CL, the fact that both MAR systems tended to reduce participants' mental demand during Experiment I may have been because the 3D information they superimposed on the real environment (as shown in Fig. 9) facilitated their users' cognitive processes: enabling inspectors to simultaneously perform several cognitive activities, such as looking, comprehending, searching, remembering, and deciding, unlike with paper-based inspection. Also, temporal stress in Experiment I was probably less for the members of the two MAR groups than for the traditional-inspection group, because the former two sets of participants did not need to perform time-consuming gaze shifts between paper drawings and the real environment (Polvi et al., 2018). Thus, MAR's 3D superimpositions on the real environment could be said to have lowered inspector effort physically as well as mentally. And unsurprisingly, our Experiment II results confirmed that performing the same tasks in a realistically simulated hazardous construction environment increased the cognitive demands on all three groups.

In terms of performance, Experiment I established that the two MAR systems' superimposed 3D rebar models increased the participants' performance when it came to detecting errors in the numbers of rebars or their spacing. Also, MARS has the potential to allow its user to more focus on the task by reducing the number of necessary gaze shits between the real and augmented environment, and thus user's performance is expected to increase (Polvi et al., 2018). However, the reduction in the number of these shifts can vary according to the AR-assisted display system. However, both MAR systems we tested also appeared to have some negative impacts on inspection performance. For example, neither could provide clear depth information regarding rebar placement, due to perception issues, and this resulted in significantly lower performance by the MAR groups (as compared to the paper-inspection group) when it came to finding side-cover, bottom-cover, and tying/support errors. On the other hand, participants equipped with either version of MAR were able to complete their inspection tasks more quickly than those who were not, with the tablet group finishing quickest, probably thanks to their devices' relatively large field of view, as compared to HoloLens.

Our Experiment II results, meanwhile, confirmed that the MAR-assisted groups' TP decreased slightly more than the traditional-inspection group's did when all three groups were placed in a more realistic construction environment. In particular, task-completion time increased significantly for the HoloLens group, probably implying that HoloLens's relatively small field of view made task performance more time-consuming when the environment was more complex, distracting, and potentially hazardous.

The superimposed 3D rebar models shown in both head-mounted and tablet-based MAR appeared to help their users to understand the inspection task itself. However, both had disadvantages, relative to paper-based inspection, in terms of SA: which was observed to be lower for both groups of MAR users across all three dimensions of the SART. First, both MAR systems, but especially HoloLens, appeared to provide their users with less understanding of their surroundings (U), probably because both restrict the field of view. Generally, human have a horizontal field of view of 104 degrees to 94 degrees for each eye (over 180 degrees approximately) (Knapp and Loomis, 2004), while tablet and HoloLens have a relatively small field of view. HoloLens users, in particular, tend to keep their gaze constantly on the AR environment, making it difficult for them to fully understand their surroundings or to use their cognitive resources (i.e., arousal, concentration, attention, and mental capacity) appropriately, other than on the task at hand. Considering that inspectors on construction worksites must perform several cognitive activities simultaneously – looking, comprehending, searching, remembering, and deciding – they are generally required to achieve full understandings of their surroundings over a very short period. Our experimental results confirm that equipping inspectors with MAR, and especially head-mounted MAR, is likely to be counterproductive, as our participants in the paper-based group were more fully aware of small changes in the background environment than their MAR-assisted counterparts. In short, MAR use by AEC-industry inspectors could reasonably be expected to increase potential worksite-safety issues, in particular, due to the restrictions these devices place on their wearers' fields of view, which tend to focus their attention more narrowly on their tasks than natural human vision would, and thus render them less alert to changes and potential changes in their immediate environment.

Despite these important findings, there may be some limitations of this study. First, even though we obtained the statistically significant results from the experimental sessions, the relatively small number of participants may lead to the generalizability issues of the findings due to the human variability. For example, the task performance when using new technologies such as MAR devices could be highly affected by the user's technology acceptance or previous experience on using them (Olsson et al., 2012). To minimize the issue, comprehensive pre-training sessions were provided to participants, but the possibility of participants' different learning abilities still may remain. Also, the individual difference in participants' cognitive ability level (i.e., finding errors in rebar placement) was not fully controlled, which may lead to misinterpretation of the results. So, further studies would be needed to provide strong generalizability by considering other human variability issues in the future. Also, the self-assessment survey could suffer from potential bias in response. Participants might obtain different interpretations of questions and respond in a certain way irrespective of the content of the questions, which is known as acquiescence response bias (Kam and Meyer, 2015). More objective measures for cognitive load and situational awareness may need to be investigated. Recently, measurement techniques using sensor data such as eye-tracking or electroencephalogram (EEG) signals have been tested for measuring cognitive workload, showing the potential as objective assessment (Borys et al., 2017). Lastly, while our second laboratory experiment tried to simulate a real construction-site inspection experience as closely as possible, the complexity and uncertainty of an actual construction site are very difficult to replicate. During real inspection tasks at construction sites, the cognitive demands on workers may be even higher than reported above, leading to lowering the MAR user's task performance. By the same token, in any complex construction environment, workers need to use more cognitive resources to observe actual and possible environmental changes at construction sites. Therefore, future research should confirm the validity of the above results through a field experiment, as well as with a wider variety of MAR systems.

CONCLUSIONS

In the AEC industry, MAR is widely considered to support its users' cognitive capability via the superimposed information it provides. However, such information may lead to cognitive overload and thus could adversely effects on the performance of tasks. Also, the limited user's field of view that comes with MAR use could limit his/her ability to notice events in their surroundings. Therefore, this study compared the impact of two distinct types of MAR (i.e., handheld and head-mounted systems) on construction professionals' CL, TP, and SA, relative both to each other and to paper-based techniques. While the rebar-framework design information provided via a superimposed virtual rebar model in MAR-assisted inspection appeared to decrease the inspectors' CL associated with the information-seeking (e.g., the number of rebars required; proper spacing) and processing (e.g., identifying missing or superfluous rebars in the actual rebar framework), it negatively impacted their performance in dangerous surroundings. The head-mounted MAR device we used, in particular, decreased its users' understanding of the surrounding environment and increased their inspection-task completion times, as compared not only to paper-based inspection but also to its tablet-based counterpart. As such, the key contribution of this research is that both of the main existing modalities of MAR-based inspection influence CL, TP and SA – for the most part, negatively.

Despite the aforementioned limitations of this study, several theoretical and practical implications can be derived from the results. The findings of both our experiments can contribute to the body of knowledge that a given information-presentation format can influence construction practitioners' cognitive workload and performance during MAR-supported tasks. Also, the findings of the research could provide a better understanding of MAR cognitive issues. In addition, the findings of our research would guide the design and usage of MAR systems for construction tasks, and this could possibly enhance the human cognitive functioning at construction worksites by better utilization of MAR systems.

DATA AVAILABILITY STATEMENT

All models used for this study (e.g., MAR models) and the data that support the findings (e.g., survey results) are available from the corresponding author upon reasonable request.

ACKNOWLEDGMENTS

This research study was supported by the Start-up Fund project (No. 1-ZE6Y) from the Hong Kong Polytechnic University, Hong Kong, and a grant (19CTAP-C151784-01) from Technology Advancement Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.

REFERENCES

Agarwal, S. 2016. "Review on application of augmented reality in civil engineering." In *Proc. of International Conference on Inter Disciplinary Research in Engineering and Technology*, 68-71. London: ASDF International.

Ahad, N. A., Yin, T. S., Othman, A. R., and Yaacob, C. R. 2011. "Sensitivity of normality tests to non-normal data." *Sains Malays.* 40 (6): 637-641.

Alsafouri, S., and Ayer, S. K. 2019. "Mobile Augmented Reality to Influence Design and Constructability Review Sessions." *J Architect Eng.* 25 (3): 1-11. https://doi.org/10.1061/(ASCE)AE.1943-5568.0000362.

Ammari, K. E., and Hammad, A. 2014. "Collaborative BIM-based markerless mixed reality framework for facilities maintenance." In *Proc. of International Conference on Computing in Civil and Building Engineering*, 657-664. Florida: ASCE.

Bae, H., Golparvar-Fard, M., and White, J. 2013. "High-precision vision-based mobile augmented reality system for context-aware architectural, engineering, construction and facility management (AEC/FM) applications." *Vis. in Eng.* 1 (1): 1-13. https://doi.org/10.1186/2213-7459-1-3.

Baumeister, J., Seung Youb, S., Elsayed Jillian Dorrian, N. A. M., Webb, D. P., Walsh, J. A., Simon, T. M., Irlitti, A., Smith, R. T., Kohler, M., and Thomas, B. H. (2017). "Cognitive cost of using augmented reality displays." *IEEE T Vis Comput Gr,* 23 (11), 2378-2388. https://doi.org/10.1109/TVCG.2017.2735098

Bhandari, S., Hallowell, M. R., Van Boven, L., Golparvar-Fard, M., Gruber, J., and Welker, K. M. 2018. "Using Augmented Virtuality to Understand the Situational Awareness Model." In *Proc. of Construction Research Congress*, 105-115. Louisiana.

Bhandary, S., Lipps, J., Winfield, S., Abdel-Rasoul, M., and Stoicea, N. 2016. "NASA Task Load Index Scale to Evaluate the Cognitive Workload during Cardiac Anesthesia Based

Simulation Scenarios." *Int J Anesth Res.* 4 (8): 300-304. http://dx.doi.org/10.19070/2332-2780-1600063.

Borys, M., Tokovarov, M., Wawrzyk, M., Wesolowska, K., Plechawska-Wojcik, M., Dmytruk, R., and Kaczorowska, M. (2017). "An analysis of eye-tracking and electroencephalography data for cognitive load measurement during arithmetic tasks." In *Proc., 10th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*, 287-292. Bucharest: IEEE

Chalhoub, J., Ayer, S. K., and Applications. 2019. "Exploring the performance of an augmented reality application for construction layout tasks." *Multimed Tools Appl*: 1-24. https://doi.org/10.1007/s11042-019-08063-5.

Curtin, A., and Ayaz, H. 2017. "Cognitive Considerations in Auditory User Interfaces: Neuroergonomic Evaluation of Synthetic Speech Comprehension " In *Proc. of 14th International Conference, Engineering psychology and cognitive ergonomics (EPCE)*. Vancouver: Springer.

Dadi, G. B., Goodrum, P. M., Taylor, T. R., and Carswell, C. M. 2014. "Cognitive workload demands using 2D and 3D spatial engineering information formats." *J Constr Eng M ASCE.* 140 (5): 1-8. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000827.

Darko, A., and Chan, A. P. C. 2018. "Strategies to promote green building technologies adoption in developing countries: The case of Ghana." *Build Environ.* 130: 74-84. https://doi.org/10.1016/j.buildenv.2017.12.022.

Demetriou, C., Ozer, B. U., and Essau, C. A. (2015). Self-report questionnaires. The encyclopedia of clinical psychology. 1-6. Hoboken: John Wiley & Sons, Inc.

Dey, A., Billinghurst, M., Lindeman, R. W., and Swan, J. 2018. "A systematic review of 10 years of augmented reality usability studies: 2005 to 2014." *Frontiers in Robotics and AI.* 5: 1-28. https://doi.org/10.3389/frobt.2018.00037.

Dodd-McCue, D., and Tartaglia, A. J. C. T. (2010). "Self-report response bias: Learning how to live with its diagnosis in chaplaincy research." *Chaplaincy Today,* 26 (1): 2-8. https://doi.org/10.1080/10999183.2010.10767394.

Doswell, J. T., and Skinner, A. (2014). "Augmenting human cognition with adaptive augmented reality." In *Proc. of the International Conference on Augmented Cognition*, 104-113. Heraklion: Springer.

Dunston, P. S. J. A. i. C. 2009. "Evaluation of augmented reality in steel column inspection." *Autom. Constr.* 18 (2): 118-129. https://doi.org/10.1016/j.autcon.2008.05.007.

Endsley, M. R. 1988a. "Design and evaluation for situation awareness enhancement." In *Proc. of Human Factors Society Annual Meeting*, 97-101. Los Angeles, CA: SAGE Publications.

Endsley, M. R. 1988b. "Situation awareness global assessment technique (SAGAT)." In *Proc. of National Aerospace and Electronics Conference*, 789-795. IEEE.

Endsley, M. R., and Garland, D. J. 2000. *Situation awareness analysis and measurement*. New Jersey: Lawrence Erlbaum Associates.

Endsley, M. R., Selcon, S. J., Hardiman, T. D., and Croft, D. G. 1998. "A comparative analysis of SAGAT and SART for evaluations of situation awareness." In *Proc. of the Human Factors and Ergonomics Society Annual Meeting*, 82-86. Los Angeles, CA: SAGE Publications.

Hart, S. G. 2006. "NASA-task load index (NASA-TLX); 20 years later." In *Proc. of the human factors and ergonomics society annual meeting*, 904-908. Los Angeles, CA: Sage Publications.

Hasanzadeh, S., Esmaeili, B., and 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." *J Constr Eng M ASCE.* 144 (7): 1-18. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001516.

Heinzel, A., Azhar, S., and Nadeem, A. 2017. "Uses of Augmented Reality Technology during Construction Phase." In *Proc. of the Ninth International Conference on Construction in the 21st Century (CITC-9)*, 866-873. Dubai, United Arab Emirates.

Hou, L., Wang, X., and Truijens, M. 2013. "Using augmented reality to facilitate piping assembly: an experiment-based evaluation." *J. Comput. Civ. Eng.* 29 (1): 1-12. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000344.

Irizarry, J., Gheisari, M., Williams, G., and Walker, B. N. 2013. "InfoSPOT: A mobile Augmented Reality method for accessing building information through a situation awareness approach." *Autom. Constr.* 33: 11-23. https://doi.org/10.1016/j.autcon.2012.09.002.

Izkara, J. L., Pérez, J., Basogain, X., and Borro, D. 2007. "Mobile augmented reality, an advanced tool for the construction sector." In *Proc. of Conf. on Bringing Information and Communication Technology (ICT) Knowledge to Work*, 190-202. Rotterdam: Int. Council for Research and Innovation in Building and Construction.

Joseph, S. (2013). *Measuring cognitive load: A comparison of self-report and physiological methods.* Ph.D thesis, Arizona State University, Tempe, AZ. Retrieved from https://repository.asu.edu/attachments/110550/content/SchinkJoseph_asu_0010E_12971.pdf.

Kalawsky, R., Hill, K., Stedmon, A. W., Cook, C., and Young, A. 2000. "Experimental research into human cognitive processing in an augmented reality environment for embedded training systems." *Virtual Real-London.* 5 (1): 39-46. https://doi.org/10.1007/BF01418975.

Kam, C. C. S., and Meyer, J. P. (2015). "How careless responding and acquiescence response bias can influence construct dimensionality: The case of job satisfaction." *Organ. Res. Methods,* 18 (3): 512-541. https://doi.org/10.1177/1094428115571894.

Khawaja, M. A., Chen, F., and Marcus, N. 2014. "Measuring cognitive load using linguistic features: implications for usability evaluation and adaptive interaction design." *Int J Hum Comput Interact.* 30 (5): 343-368. https://doi.org/10.1080/10447318.2013.860579.

Kim, Y.-T., Lee, J.-Y., Lee, S.-H., and Choi, J.-S. 2013. "Construction and inspection management system using mobile augmented reality." In *Proc. of 19th Korea-Japan Joint Workshop on Frontiers of Computer Vision*, 93-96. IEEE.

Knapp, J. M., and Loomis, J. M. 2004. "Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments." *Presence-Teleop Virt.* 13 (5): 572-577. https://doi.org/10.1162/1054746042545238.

Kopsida, M., and Brilakis, I. 2016. "BIM registration methods for mobile augmented reality-based inspection." In *Proc. of 11th European Conference on Product and Process Modelling (ECPPM)*, 201. Limassol, Cyprus: CRC Press.

Kruijff, E., Swan, J. E., and Feiner, S. 2010. "Perceptual issues in augmented reality revisited." In *Proc. of International Symposium on Mixed and Augmented Reality*, 3-12. Seoul: IEEE.

Krumpal, I. (2013). "Determinants of social desirability bias in sensitive surveys: a literature review." *Qual Quant*, 47 (4): 2025-2047. https://doi.org/10.1007/s11135-011-9640-9.

Kwon, O.-S., Park, C.-S., and Lim, C.-R. 2014. "A defect management system for reinforced concrete work utilizing BIM, image-matching and augmented reality." *Autom. Constr. 46*: 74-81. https://doi.org/10.1016/j.autcon.2014.05.005.

Lamsal, B., and Kunichika, K. 2019. "Development of an AR System for the Advancement of the Tasks in the Construction sites." In *Proc. of Creative Construction Conference*, 830-835. Budapest, Hungary.

Lee, B., Chung, K., and Kim, S.-H. 2018. "Interruption Cost Evaluation by Cognitive Workload and Task Performance in Interruption Coordination Modes for Human–Computer Interaction Tasks." *Appl. Sci.* 8 (10): 1-20. https://doi.org/10.3390/app8101780.

Lei, H., Xiangyu, W., Leonhard, B., and Peter, L. 2013. "Using animated augmented reality to cognitively guide assembly." *J. Comput. Civ. Eng.* 27 (5): 439-451. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000184.

Li, N., and Duh, H. B.-L. (2013). Cognitive issues in mobile augmented reality: An embodied perspective. 109-135. New York: Springer.

Lindblom, J., and Thorvald, P. 2014. "Towards a framework for reducing cognitive load in manufacturing personnel." In *Proc. of the 5th International Conference on Applied Human Factors and Ergonomics (AHFE)*, 6267-6278. Kraków.

Longo, L. 2018. "Experienced mental workload, perception of usability, their interaction and impact on task performance." *PloS one.* 13 (8): 1-36. https://doi.org/10.1371/journal.pone.0199661.

Lukosch, S., Lukosch, H., Datcu, D., and Cidota, M. 2015. "Providing information on the spot: Using augmented reality for situational awareness in the security domain." *Comput Supp Coop W J.* 24 (6): 613-664. https://doi.org/10.1007/s10606-015-9235-4.

Lyell, D., Magrabi, F., and Coiera, E. 2018. "The effect of cognitive load and task complexity on automation bias in electronic prescribing." *Human factors.* 60 (7): 1008-1021. https://doi.org/10.1177/0018720818781224.

Matthews, M. D., and Beal, S. A. 2002. *Assessing situation awareness in field training exercises, Reseasrch Report 1795*. Retrieved from U.S. Army Research Institute for the Behavioural and Social Sciences.:

Matthews, M. D., Martinez, S. G., Eid, J., Johnsen, B. H., and Boe, O. C. 2005. "A comparison of observer and incumbent ratings of situation awareness." In *Proc. of the Human*

Factors and Ergonomics Society Annual Meeting, 548-551. Los Angeles, CA: SAGE Publications.

McGuinness, B., and Foy, L. 2000. "A subjective measure of SA: the Crew Awareness Rating Scale (CARS)." In *Proc. of Human Performance, Situation Awareness, and Automation Conference*, 286-291. Savannah, Georgia.

McKight, P. E., and Najab, J. 2010. *Kruskal-wallis test. The corsini encyclopedia of psychology*. New York: Wiley.

Mendel, J., and Pak, R. 2009. "The effect of Interface Consistency and Cognitive Load on user performance in an information search task." In *Proc. of Human Factors and Ergonomics Society Annual Meeting*, 1684-1688. Los Angeles, CA: SAGE Publications.

Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., and Keshri, A. 2019. "Descriptive statistics and normality tests for statistical data." *Ann Card Anaesth.* 22 (1): 67-72. http://dx.doi.org/10.4103/aca.ACA_157_18.

Naderpour, M., Lu, J., and Zhang, G. 2016. "A safety-critical decision support system evaluation using situation awareness and workload measures." *Reliab Eng Syst Safe.* 150: 147-159. https://doi.org/10.1016/j.ress.2016.01.024.

Olsen, D., Kim, J., and Taylor, J. M. 2019. "Using Augmented Reality for Masonry and Concrete Embed Coordination." In *Proc. of Creative Construction Conference*, 906-913. Budapest.

Olsson, T., Kärkkäinen, T., Lagerstam, E., and Ventä-Olkkonen, L. J. J. o. A. I. (2012). "User evaluation of mobile augmented reality scenarios." *J Amb Intel Smart En*, 4 (1), 29-47. https://doi.org/10.3233/AIS-2011-0127.

Paas, F., Renkl, A., and Sweller, J. 2003. "Cognitive load theory and instructional design: Recent developments." *Educ. Psychol.* 38 (1): 1-4. https://doi.org/10.1207/S15326985EP3801_1.

Paas, F., Renkl, A., and Sweller, J. 2004. "Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture." *Instr. Sci.* 32 (1-2): 1-8. https://doi.org/10.1023/B:TRUC.0000021806.17516.d0.

Polvi, J., Taketomi, T., Moteki, A., Yoshitake, T., Fukuoka, T., Yamamoto, G., Sandor, C., and Kato, H. 2018. "Handheld guides in inspection tasks: augmented reality versus picture." *IEEE T Vis Comput Gr.* 24 (7): 2118-2128. https://doi.org/10.1109/TVCG.2017.2709746.

Rankohi, S., and Waugh, L. 2013. "Review and analysis of augmented reality literature for construction industry." *Vis. in Eng.* 1 (1): 1-9. https://doi.org/10.1186/2213-7459-1-9.

Salmon, P., Stanton, N., Walker, G., and Green, D. 2006. "Situation awareness measurement: A review of applicability for C4i environments." *Appl Ergon.* 37 (2): 225-238. https://doi.org/10.1016/j.apergo.2005.02.001.

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., and Young, M. 2009. "Measuring Situation Awareness in complex systems: Comparison of measures study." *Int J Ind Ergonom.* 39 (3): 490-500. https://doi.org/10.1016/j.ergon.2008.10.010.

Stanton, N. A., Salmon, P. M., Rafferty, L. A., Walker, G. H., Baber, C., and Jenkins, D. P. 2005. *Human factors methods: a practical guide for engineering and design*. Aldershot: Ashgate Pub. Co.

Taylor, R. 1990. "Situation Awareness Rating Technique: the development of a tool for aircrew system design." In *Proc. of Situation Awareness in Aerospace Operation (AGARD)*, 3/1-3/17. Neuilly Sur Seine: NATO-AGARD.

Wang, X. 2007. "Using augmented reality to plan virtual construction worksite." *Int J Adv Robot Syst.* 4 (4): 501-512. https://doi.org/10.5772/5677.

Woodward, C., and Hakkarainen, M. 2011. "Mobile augmented reality system for construction site visualization." In *Proc. of Int. Symp. on Mixed and Augmented Reality (ISMAR)*, 1-6. New York: ISMAR.

Zaher, M., Greenwood, D., and Marzouk, M. 2018. "Mobile augmented reality applications for construction projects." *Constr. Innov.* 18 (2): 152-166. https://doi.org/10.1108/CI-02-2017-0013.

Zhou, Y., Luo, H., and Yang, Y. 2017. "Implementation of augmented reality for segment displacement inspection during tunneling construction." *Autom. Constr.* 82: 112-121. https://doi.org/10.1016/j.autcon.2017.02.007.

FIGURE CAPTION LIST

Table 1. The Five NASA-TLX Questions Used for Measuring Cognitive Load (Hart, 2006)

Dimension	Question
Mental Demand	How mentally demanding was the task?
Temporal Demand	How temporally demanding was the task?
Performance	How successful were you in accomplishing what you were asked to do?
Effort	How hard did you have to work to achieve your level of performance?
Frustration	How insecure, discouraged, irritated, or stressed were you during the task?

Table 2. Items for Measuring Situational Awareness (Taylor, 1990)

Domain	Items	Questions
Understanding (U)	Information Quantity (1)	How much information about your surroundings did you take in?
	Information Quality (2)	How well did you understand/comprehend the information about your surroundings that you took in?
	Familiarity (3)	How familiar with your surroundings did you become during the task?
Attentional Demand (D)	Instability (4)	How much was the situation in your surroundings changing during the experimental session?
	Variability (5)	Were a number of different factors in the surrounding environment changing?

	Complexity (6)	How complex was the surrounding situation?
Attentional Supply (S)	Arousal (7)	How alert were you to observing the surrounding situation?
	Concentration (8)	How much were you concentrating on your surroundings?
	Division of Attention (9)	What proportion of your attention was devoted to your surroundings, as opposed to your inspection task?
	Spare Mental Capacity (10)	How much mental capacity did you have to spare for your surroundings?

Table 3. Average Number of Errors Correctly Identified, by Inspection-modality Group

Rebar errors	Mediums	No. of Errors Placed	Experiment I Mean (SD)	Kruskal-Wallis H	*p*
Spacing between bars	Paper	5	2.73 (1.43)	6.54	0.03**
	Tablet		2.06 (0.88)		
	HoloLens		3.00 (0.75)		
Missing rebars	Paper	2	1.60 (0.63)	0.52	0.77*
	Tablet		1.53 (0.45)		
	HoloLens		1.46 (0.63)		
Extra rebars	Paper	3	1.66 (0.48)	3.80	0.14*
	Tablet		2.13 (0.99)		
	HoloLens		2.00 (0.84)		
Incorrect side-cover spacing	Paper	2	1.66 (0.61)	18.55	0.00**
	Tablet		0.44 (0.83)		
	HoloLens		0.40 (0.63)		
Incorrect bottom-cover spacing	Paper	2	0.60 (0.82)	7.96	0.01**
	Tablet		0.06 (0.25)		
	HoloLens		0.06 (0.25)		
Incorrect number of anchorage bars	Paper	2	2.00 (0.00)	6.27	0.04**
	Tablet		1.66 (0.72)		
	HoloLens		2.00 (0.00)		
	Paper	2	1.73 (0.73)	15.61	0.00**
	Tablet		1.00 (0.75)		

Incorrect length of anchorage bars	HoloLens		0.46 (0.74)		
Bars improperly tied and supported	Paper	2	1.53 (0.74)	28.82	0.00**
	Tablet		0.20 (0.56)		
	HoloLens		0.00 (0.00)		
Total number of errors	Paper	20	13.51 (5.44)	19.61	0.00**
	Tablet	20	9.08 (5.43)		
	HoloLens	20	9.42 (4.29)		

Note. *=No significant difference (p>0.05); **=Significant difference (p<0.05).

Table 4. Average Number of Errors Correctly Identified, Experiments I and II

Mediums	Number of Errors Placed	Total Errors Identified in Experiment I, Cumulative Mean (SD)	Total Errors Identified in Experiment II, Cumulative Mean (SD)	Kruskal-Wallis H	*p*
Paper	20	13.51 (5.44)	14.58 (4.70)	2.94	0.08*
Tablet	20	9.08 (5.43)	8.53 (4.12)	0.69	0.40*
HoloLens	20	9.42 (4.29)	8.72 (4.98)	0.25	0.61*

Note. *=No significant difference (p>0.05); **=Significant difference (p<0.05).

Table 5. Situation Awareness Rating Technique Scores

SART Item	Modality			Kruskal-Wallis H	*p*
	Paper Mean (SD)	Tablet Mean (SD)	HoloLens Mean (SD)		
Information Quantity (1)	3.40 (0.91)	3.33 (1.29)	2.93 (0.79)	2.10	0.34*
Information Quality (2)	3.0 (0.84)	3.13 (1.35)	3.00 (0.75)	0.00	0.99*
Familiarity (3)	3.40 (0.98)	3.07 (1.38)	2.93 (0.35)	2.06	0.35*
Instability (4)	3.46 (0.92)	3.26 (0.79)	3.24 (0.70)	0.68	0.70*
Variability (5)	3.40 (0.73)	2.86 (0.91)	2.85 (0.83)	4.18	0.12*

Complexity (6)	3.46 (0.64)	3.20 (0.67)	3.06 (0.79)	2.13	0.31*
Arousal (7)	3.33 (0.74)	2.66 (1.23)	2.86 (0.91)	4.75	0.09**
Concentration (8)	3.20 (0.94)	2.73 (1.16)	3.00 (0.84)	1.30	0.52*
Division of attention (9)	3.40 (1.05)	2.80 (1.08)	2.66 (0.61)	4.16	0.12*
Spare mental capacity (10)	3.46 (0.91)	3.06 (0.79)	2.73 (0.79)	2.10	0.34*
Understanding (U)	9.8 (2.73)	9.53 (4.02)	8.86 (1.89)	2.33	0.31*
Attentional Demand (D)	10.31 (2.29)	9.26 (2.37)	9.18 (2.32)	7.02	0.03**
Attentional supply (S)	13.39 (3.64)	11.26 (4.26)	11.25 (3.15)	12.70	0.00**
SART=U-[D-S]	12.88 (4.08)	11.53 (5.91)	10.93 (2.72)	1.23	0.53*

Note. *=No significant difference ($p>0.05$); **=Significant difference ($p<0.05$).