Effect of Augmented Reality User Interface on Task Performance, Cognitive Load, and Situational Awareness in Human-Robot Collaboration

Apostolos Kalatzis, Vishnunarayan Girishan Prabhu, Laura Stanley, and Mike P. Wittie

***Abstract*— Augmented Reality (AR) enables the transmission of intent using the physical area in which humans and robots interact as a shared canvas. Studies exploring AR for human-robot collaboration have reported mixed findings on the relationship between cognitive workload and task performance. In this study, we developed an AR user interface (UI) that guides the user to perform a pick-and-place task while collaborating with a robot. A repeated measures mixed-methods study with sixteen participants demonstrated that AR UI significantly impacted task performance, where users traveled longer distances at slower speeds to pick-and-place objects than the control group. Additionally, UI significantly impacted the cognitive load, where participants reported higher NASA-TLX scores while using AR UI. Finally, users reported significantly lower situational awareness and low usability scores while using AR UI. Our findings suggest that the AR UI negatively impacts human-robot collaboration, calling for further investigation.**

Human-robot collaboration, cognitive workload, human factors, mixed reality

I. Introduction

The fourth industrial revolution, or Industry 4.0, primarily aimed to adapt and utilize collaborative robots for human-robot collaboration to improve safety, accuracy, performance, and reduce workload [1], [2]. Another aspect of Industry 4.0 is to develop new forms of human-machine interaction, such as touch interfaces, haptic systems, virtual reality (VR) and augmented reality (AR) systems, etc., to reduce errors and improve safety and efficiency [3]. While Industry 4.0 primarily focuses on increasing interconnectivity and smart automation to achieve optimum performance, productivity and enhance efficiency, Industry 5.0 focuses on refining the interaction between humans, machines, and robots by developing human-centered design solutions where humans and robots can collaborate to enable personalized autonomous actions and solutions through enterprise social networks [4]–[6].

A user interface (UI) plays a critical role as the communication gateway between users and robots, allowing them to perform various actions efficiently and effectively to achieve goals [7], [8]. UIs range from simple 2-D displays to complex state-of-the-art 3-D displays (e.g., VR, AR, etc.). However, irrespective of the type of UI, some characteristics and features, including consistency, feedback, reversal of action, support internal locus of control, reduced short-term memory load, etc., are vital for a good interface design. Although these are crucial factors for efficient functioning and human-robot collaboration, the short-term memory load and cognitive load that interface asserts on humans should be given special attention when considering state-of-the-art head-mounted extended reality (XR) devices, including augmented reality (AR) and virtual reality (VR) headsets [9], [10]. We refer to the broad category of state-of-the-art XR devices as any HMD devices that are semi- or fully immersive, including Oculus, HTC Vive, HoloLens, etc. Studies investigating the use of XR technologies for complex human-robot collaborative tasks and daily tasks have observed mixed findings where few observed that the users experience a higher cognitive workload and vice versa [11]–[17]. The cognitive workload cannot be directly attributed to the XR technology as various factors, including task complexity, duration of the task, interface design, and device, plays a significant role [18]–[20]. Primary reasons attributed to an increased cognitive workload while using XR technology for human-robot collaboration include cumbersome technology, time of use, the field of view, lack of feedback, delay in processing feedback, and task complexity [21], [22]. Hence, researchers and practitioners should consider these factors and focus on improving the interface design (feedback, response) and other parameters (task complexity, time of use) while using head-mounted devices, including AR and VR headsets, for human-robot collaboration.

Situational awareness is another critical factor to consider while developing interfaces for human-robot collaboration. Situational awareness or Situation Awareness (SA) is the ability to perceive and comprehend what is happening in the environment around you at any given time or space and predict what can happen in the future [23]. Prior studies have validated that SA significantly impacts decision-making and is affected by the cognitive load and resources available while making decisions [24], [25]. Researchers have reported a lack of SA as a primary cause leading to errors, failures, and fatalities in complex systems such as aviation, manufacturing, and mining [26]–[29]. While chances for lapses in SA are common in complex environments, their likelihood in environments involving higher automation (human-robot collaboration, autonomous driving, etc.), along with assistive technology, is not well-researched [30], [31]. The consensus is that increasing automation and providing assistive technology during human-robot collaboration can offload the burden on the user and reduce their cognitive load leading to a better SA. However, few studies have investigated how AR interfaces impact SA during human-robot collaboration.

Apostolos Kalatzis, Laura Stanley, and Mike P. Wittie are with the Gianforte School of Computing, Montana State University, Bozeman, MT, 59717 USA

Vishnunarayan Girishan Prabhu is with the Systems Engineering and Engineering Management, University of North Carolina at Charlotte, NC, 28223 USA.

We developed an AR assistive interface for human-robot collaboration to address these gaps. We compared its impact on users' cognitive load, task performance, and situational awareness while performing a collaborative pick-and-place task to a control group performing the same task. Using validated questionnaires, physiological biomarkers, and other objective data to capture cognitive load, task performance, and situational awareness, we hypothesize that AR assistive interface for human-robot collaboration will improve task performance and situational awareness and reduce the cognitive workload during a human-robot collaborative pick-and-place task.

II. Relative Work

A. Augmented reality and Human-Robot Interaction

In human-robot collaboration, robot motion inference or otherwise, the gulf of evaluation refers to the lack of ability of a user to understand and assess a system [32], [33]. For example, a user might believe that a robot perceives and tracks the user and the surrounding environment, but this is untrue. This results from robots' inability to communicate cues and information between the robot and the user. The information exchange is usually based on two dimensions: the communication medium and the communication format. Currently, the two commonly manifested media in human-robot collaboration are: 2-D displays and HMDs (VR, AR, etc.) [34]–[37]. However, 2-D displays might not be able to communicate complex paths and motions transparently, leading to a focus switch from the real-world environment to the communication tool leading to an overall delay in communication [35], [38]. Therefore, the design of intuitive user interfaces is critical to unfold the full potential of human skills and provide more efficient human-robot interaction [7]. Intuitive AR media presented as graphical user interfaces well mapped to specific human-robot collaboration tasks can allow users to focus on task goals [39]. Recent research has investigated using AR to guide the user to complete a task. Text instructions are usually provided to navigate the user in human-robot collaborative tasks similar to ones performed on industrial lines [11]–[16]. These texts contain information regarding the direction that should be taken to complete a task, the rotation direction or how much rotation is required, and information regarding the final destination. Another form of guidance is the representation of the holographic items needed in a task [11]–[17]. These representations guide the location of the items to complete an assembly or the type of item needed. Finally, virtual guidance in the form of arrows has been proposed to help the users complete a task [11]. These arrows provide information about the next move at each stage of the task [11]. Another recent study developed an AR interface that provides geometrical path planning, spatial mapping, and virtual models with arrows for completing a virtual pick-and-place task [40]. A few other recent studies investigating pick-and-place tasks have focused on robot programming using AR interfaces with the help of augmented trajectories. These studies have reported AR to be easier compared to using a keyboard for navigation [41], [42]. Prior studies have utilized AR in human-robot collaboration primarily for various assembly tasks as a gateway for collaboration in assembling various components and in some pick-and-place tasks to comprehend the robot programming. While these are critical in advancing our knowledge to develop collaborative robots, this study focuses on evaluating an AR-based UI that guides the user to complete a task.

B. Cognitive Workload and Human-Robot Interaction

In an environment involving complex human-machine systems and other interconnected intelligent machines where the role of the human is critical, monitoring users' cognitive workload is crucial for ensuring efficiency and safety [43], [44]. Pick-and-place is among the most common tasks in manufacturing and assembly lines where industrial robots are used extensively [45]–[47]. While picking randomly positioned parts from a surface and placing them into another place is logically a simple task for a human but could be physically challenging. Whereas for robots, this same task could be logically and logistically challenging [48]. A major difficulty is locating and gripping components in the correct orientation [48]. However, robots can be controlled manually, where a good level of precision can be achieved using remote joysticks. To accomplish this, humans and robots should form a dynamic system working together to achieve a common goal. One way to achieve this collaborative, dynamic system is by incorporating intuitive media presented as graphical user interfaces mapped to specific human-robot interaction tasks can allow users to focus on task goals and lead to more efficient human-robot interaction [7]. Specifically, AR has immense potential as assistive technology in human-robot collaboration by providing users with helpful information when needed, and previous research has shown that guidance and holographic information can lead to better performance results and augment the human-robot collaboration [11]–[17]. However, the effect of AR on cognitive workload and task performance is unclear as researchers have shown both positive and negative impacts of using AR to present information to a user. Some studies suggest that AR positively affects cognitive workload and task performance [39], [49]–[51]. Whereas some studies show evidence of cognitive overload using AR. But it should be noted that the overload is primarily attributed to ergonomic and usability issues and, more specifically, to the prolonged use of technology, the field of view, and lag in providing the information to the user [18], [52]–[54]. Additionally, one study that reported AR induces lower cognitive workload as compared to a 2-D interface for a pick-and-place task used gestures in the AR interface and keyboards and mice in the 2-D interface to perform the task, which is a significant confounder [41].

Although the findings from these studies offer important insights, one primary limitation is that all these studies relied solely on subjective feedback (i.e., NASA TLX, NASA RTLX, surgical TLX) to measure cognitive workload. While subjective perceptions of cognitive workload are valuable and collected using validated tools, it also leads to many challenges since they cannot be objectively compared. Moreover, in most studies, operators often report their cognitive workload at the end of the task, which fails to capture the cognitive load in real time while performing a task. This calls for utilizing objective methods that can monitor cognitive load unobtrusively while performing a task. One such validated method is eye-tracking, which can capture cognitive workload continuously. One of the most popular and validated objective methods to seamlessly monitor a user's cognitive workload is pupillometry, and studies have used various metrics, including pupil diameter, blinks, etc., to quantify the change in cognitive workload [55]–[62]. Researchers have reported a positive correlation between pupil diameter, task difficulty, and cognitive load [56], [63], suggesting that cognitive workload corresponds to pupil dilation [57], [64], [65]. The increase in pupil diameter while performing a cognitively demanding task is associated with the mental effort required to complete a task [66]. However, when the limits of processing information are exceeded, pupil dilation reaches the maximum, and no further increases are noticed [64]. Blinks can also provide insights into the cognitive load, where cognitively demanding tasks lead to delays in blinks, known as attentional blinks [60]. Studies have reported a decrease in blink rate as the complexity of the task and cognitive load increases [59], [61]. This is primarily attributed to the operator's focus on capturing all the provided information, especially in vision-demanding tasks [59]. Blink durations are also used to quantify and evaluate cognitive load, where studies have reported increased blink durations while performing cognitively challenging tasks [67], [68].

III. Methods

A. The Collaborative Task

As briefly mentioned in the introduction, pick-and-place is a common task in an assembly line. To replicate this task in a lab setting, we utilized a collaborative robot where participants controlled a UR3e (Universal Robots, DK) collaborative robot using joystick input controls to pick up objects from a surface and place these in a bin. Participants had access to six degrees of freedom but were asked to control the X, Y, and Z directions. The movement speed of the robot was kept uniform and was not controllable by the participants.

B. Augmented Reality User Interface

We developed an AR interface in HoloLens 2 to investigate the efficacy of augmented visuals in performing a pick-and-place task (Fig. 1). We designed four unique visuals to evaluate the assistive potential of AR interfaces. The visuals comprise two assistive path lines, one robot gripper tracking line, and a holographic robot gripper in the optimal position to pick up the current object. The two assistive path lines, yellow and white (Fig. 1(a)) point to the location of an object. The user must follow the yellow line, representing the optimal path to pick-and-place an object. To identify the optimal path, we developed an algorithm that detects the robot's end effector and the object's location to calculate the shortest path in the robot's cartesian plane. The white line provides information on the user's proximity to an object. This line is generated by calculating the non-cartesian distance between the robot end effector and the object location. The robot gripper tracking line, blue line (Fig. 1(b)), provides information on the path the user followed to pick-and-place the object. This line is calculated by tracking the user's trajectory moves during the pick-and-place task. The holographic robot gripper helps the users to position the gripper to the optimal location for pick up by completely overlapping the holographic gripper (Fig. 1(a), Fig. 1(b), Fig. 1(c)). These visuals are synchronized using QR codes scanned with the HoloLens 2 Spatial Perception Camera. This provides real-time 3d position coordinates that can be used to configure object location and robot hand movements. The result is a dynamic assistance interface that adds another layer of visual perception to benefit task performance. In a traditional setting, a HoloLens 2 (or other AR devices) is controlled using hand gestures or a remote. However, in the study, we developed intuitive voice commands that allow the participant to activate, deactivate, or calibrate the visuals to prevent the need for buttons or distracting inputs. For example, when the participant says "show lines" the optimal (yellow) and functional (white) line guides are set to active. A holographic robot gripper in optimal object pick-up position is shown when a participant says, "show gripper." This is beneficial as the participant is both maneuvering a joystick and focusing on the robot's movements. We hypothesize that the dynamic interface and compact visualizations overcome the common augmented reality constraint of small field-of-view (FoV) to enhance task completion and performance.

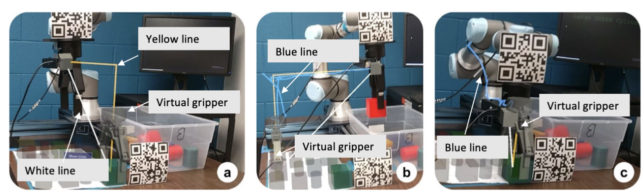

Fig. 1: User Interface

C. Participants

Participants for this study included sixteen dominant right-hand students (9M, 7F) attending a public University. All participants had an age distribution of 25.25 $\pm$ 4.31 years. Participants in the study signed a consent form and were subject to the experiment protocols approved by University's Institutional Review Board under IRBNO:2022-233. Upon consent, participants reported their experience with industrial robots, joysticks, and AR. None of the participants reported prior experience with robots and AR. Additionally, the average participant reported slight familiarity with joystick devices utilized by the participant to control the robot. Participants were compensated with a $15 Amazon gift card for their time.

D. Procedure

In this study, participants completed two sessions; in one session, the participants performed the task without the AR UI guidance (no UI AR), and in the other session utilized the developed AR UI. The order of the sessions was counterbalanced to control for order effects. Further, to mitigate the learning or order effect, participants were allowed to practice the task and use AR UI until they felt comfortable and attained equal experience. In each session, the participants picked and placed eight objects. In the UI AR session, the participants started a pick-and-place trial by saying the color and shape of the object they wanted to pick-and-place (i.e., red cube) and then the activation of the visuals and the holographic gripper using voice commands (i.e., show lines, show gripper). At the end of the trial, the participants deactivated the visuals and the holographic gripper using voice commands again. The participants did hot have to memorize any commands or the order of the objects that needed to pick up. The experimenter provided this information to the participants.

E. Dependent Variables

1) Task performance: Trajectories of each pick-and-place object were recorded at a frequency of 100 Hz. Task performance is quantified using two different measurements: (i) average speed, which is defined as the traveled trajectory length divided by the time to pick-and-place an object, and (ii) travel distance, measured by calculating the distance covered to pick-and-place each object. The distance and time to pick up the objects in the UI AR session start after the participants activate the visuals using the voice command in order to have a fair comparison between the two sessions.

2) Subjective Responses: After each session, the participants completed the situation awareness rating technique (SART) to measure situational awareness and the NASA TLX to quantify cognitive workload. SART consists of three subscales: attentional supply, attentional demand, and understanding of the task. The SART composite score was calculated as understanding-(demand-supply). NASA TLX is a multidimensional assessment tool consisting of six subjective subscales from very high to very low in terms of their: mental demand, physical demand, temporal demand, effort, performance, and frustration. In this study, we calculated the overall workload by averaging the score of the six subscales. Finally, the participants were asked to complete a system usability scale (SUS) questionnaire to understand the usability of the AR interface.

3) Objective Response: Eye-tracking: We analyzed three eye-tracking metrics based on the literature review (see Section II-B). Specifically, we collected blink rates and duration throughout the pick-and-place task to monitor and comprehend the cognitive workload. Data were collected at 200 Hz using the Pupil Labs add-on devices during each session.

4) Statistical Analysis: Statistical significance was determined through paired t-tests, or Mann-Whitney U tests on all dependent responses, with significance reported at α = 0.05. We used a parametric test for the dataset created from the study. The normality of the data for the parametric test was determined using a Q-Q plot. Equality of variances was verified using Levene's test and using boxplots. Independence was assured by our counterbalanced design. The t-tests were separately run on each task performance metric and subjective responses to test the effects of the MR UI. Separate Mann-Whitney U tests were performed on eye metrics to test the effects of the MR UI on blink count and blink duration while completing the task.

IV. Results

A. Task Performance

Table I represents the average values of the two task performance metrics (i.e., travel distance, average speed) collected in the AR UI and no AR sessions. The t-test showed a significant increase in travel distance (t=5.29, p-value$<$0.01) in picking and placing the objects using the AR UI. Additionally, a significant difference was observed in the average speed of picking and placing the objects (t=7.55, p-value$<$0.01).

B. Subjective Responses

Table II represents the average values of the six subscales of the NASA TLX questionnaire collected in the AR UI session and the no AR session. On analyzing the response collected using NASA TLX, the t-test showed a significant increase in mental demand (t=3.01, p-value $<$0.01), effort (t=3.20, p-value$<$0.01) and frustration (t=3.19, p-value$<$0.01) for the AR UI session. Additionally, performance failure perception increased significantly (t=4.46, p-value$<$0.01) after the AR UI session. Further, the AR UI session significantly increased the overall workload (t=3.67, p-value $<$0.01). However, the scores of physical demand (t=0.30, p-value=0.77) and temporal demand (t = 0.17, p-value=0.87) did not vary significantly between the two groups.

Next, we performed a t-test with the response collected using the SART questionnaire. The t-test showed a significant difference in the overall SART scores between the two groups (t=4.08, p-value$<$0.01), where participants showed low SA during the AR session. Further, to evaluate which components of the SART questionnaires varied significantly, we performed another set of t-tests with the sub-scales of SART, which included attentional demand, attentional supply, and understanding. The t-test showed a significant difference in the attentional demand (t=5.19, p-value$<$0.01), where participants reported high scores during the AR session. However, the attentional supply (t=0.27, p-value=0.79) and the understanding (t=1.60, p-value=0.13) scores did not vary significantly between the AR UI and no AR UI sessions. Finally, on analyzing the SUS score collected using the systems usability scale for comprehending the usability of the AR UI, the participants reported an average SUS score of 60.31±17.6.

C. Eye Metrics

Table III represents the average values of the eye metrics collected in the AR UI session and the no AR session. The results indicated that the blink rate (U=150, p-value=0.67) and the blink duration (U=106, p-value=0.41) scores did not differ between the AR UI and no AR UI sessions.

V. DISCUSSION

In this study, we developed an AR UI to guide users in performing a collaborative pick-and-place task. A user study was conducted to investigate the impacts of this UI on task performance, cognitive workload, and situational awareness. Additionally, a system usability score was collected to measure the usability of the developed UI.

The key takeaways of the study include:

1) The AR UI increased cognitive workload.
2) The AR UI increased travel distances to complete the pick-and-place and place task.
3) The AR UI had a negative impact on situational awareness
4) The average reported usability score was less than the average accepted SUS score when using the AR

A. The AR UI increased cognitive workload.

The perceived workload collected using the NASA TLX questionnaire indicated that the participants reported significantly higher workload levels during the AR UI session than the no AR UI. Moreover, the task performance results (i.e, travel speed, travel distance) suggest that the task performed with the assistance of the AR interface was more cognitively demanding. These results align with the findings from a few other studies where users have reported an increased cognitive load while using AR systems [11]–[17]. Considering that the robotic pick-and-place task in both sessions remained the same, one possible explanation for the increase in cognitive workload is that additional cognitive effort is needed to interpret both the real-world surroundings and the overlayed features presented in the augmented reality interface. The AR UI had a detrimental influence on the cognitive burden potentially from this dual processing leading to an increase in cognitive effort. Another potential reason that could explain the increased workload is the users' lack of familiarity with the AR systems and the associated challenges. These findings highlight the importance of considering cognitive workload when collaborative robots are designed with AR support. Additionally, to expand on these findings, future studies should explore recruiting participants with experience using VR/AR systems or consider a longitudinal study to further control the learning effect.

B. The AR UI increased travel distances to complete the pick-and-place task.

On analyzing the efficacy of the AR UI in assisting users in performing a collaborative pick-and-place task and its impact on task performance, we observed that all the participants successfully picked and placed all the objects using the AR interface. However, the results indicated that the AR UI led to longer travel distances to pick-and-place the objects compared to the control group. This finding suggests that the participants did not follow the optimal path, leading to longer travel distances. Another possible explanation is that the participants had difficulties positioning the actual robotic gripper to the virtual gripper. This mismatch led to more movement and consequently to longer travel distances and times. This negative impact on task performance aligns with the increasing cognitive load reported while using the AR UI. Prior studies have shown that cognitive workload is a significant factor influencing human performance [69]–[73]. The scientific theory supporting the impact of workload on performance is Wickens' multiple resource theory which states that humans have only a limited amount of cognitive resources to dedicate to a task, and when the task at hand exceeds the available resources it leads to inefficiency and deteriorated performance [74]. While prior studies have reported similar findings where AR UI has a negative impact on user performance, a few other studies have reported the efficacy of augmented reality tools in human-robot collaboration [12], [15], [17]. Although our results indicate that using the AR UI had a negative impact on travel distance and travel speed, it should be noted that all the participants successfully completed the task. The decreased performance can be attributed to the lack of experience and the cumbersome AR system. Our findings indicate the efficacy of AR systems to support workers during human-robot collaboration tasks, but clearly, further development, training, and better AR systems are required for actual deployment in assembly lines.

C. The AR UI had a negative impact on situational awareness

Analyzing the data collected using the SART questionnaire, we observed that the perceived SA was significantly lower among the AR UI session group than the control group. Further, comparing the subscales of SART, we observed that the attentional demand was significantly higher during the AR UI session. These findings align with the observations of cognitive load, where participants in the AR UI session reported a higher cognitive load. However, compared to the existing literature, our findings suggest some deviations regarding situational awareness. Prior studies that compared AR-based interfaces to a control group for users operating Unmanned Aerial Vehicles (UAVs), heavy vehicles, excavators, etc. reported an improved SA [75]–[78]. However, it should be noted that one study that also investigated cognitive load along with SA while using the AR device reported that cognitive load decreased while SA increased [78]. This observation is insightful because we observed similar findings but inversely, where cognitive load increased, and situational awareness decreased. These observations indicate the inverse relationship between cognitive load and situational awareness.

Finally, these findings align with the usability score reported by the participants, which was significantly lower than the consensus score of 68 for an average usable system. Based on the open-ended interview with participants after

TABLE I: Task Performance Results

Task Performance Metrics	p-value for paired t-tests	Descriptives	
		AR UI	No AR UI
Average travel distance (m)	<0.01*	1.62±0.27	1.30±0.15
Average speed (m/s)	<0.01*	0.01±0.005	0.03±0.003

TABLE II: NASA TLX Results

NASA TLX	p-value for paired t-test	Descriptives	
		AR UI	No AR UI
Mental Demand	<0.01*	56.55±21.96	33.44±21.35
Physical Demand	0.77	19.06±14.74	18.13±17.69
Temporal Demand	0.86	30.31±18.93	29.69±16.78
Performance	<0.01*	15.95±9.70	39.06±18.99
Effort	<0.01*	50.95±25.57	28.75±17.84
Frustration	<0.01*	40.88±32.69	15.94±13.19
Composite score	<0.01*	39.47±17.63	23.65±11.94

TABLE III: Eye Metrics Results

Eye Metrics	p-value for mann-whitney test	Descriptives	
		AR UI	No AR UI
Blink rate (count)	0.67	29.18±31.20	34.43±41.00
Blink duration (sec.)	0.41	0.18±0.02	0.19±0.04

both sessions, all the participants mentioned that using AR to complete the task was not easy. However, they also mentioned that it got better as they progressed through the task, suggesting that it could be a matter of training as all participants were first-time users. Additionally, the participants mentioned that they liked the interface, but the depth perception, resolution, and field of view were limited. Users suggested adding more visual cues and text in the interface could potentially help improve the usability. Finally, a few participants reported that the AR technology is "not there" yet for complex tasks. On a separate note, evaluating each question from the SUS score generated similar themes: the users found the system cumbersome, the need for guidance from a technical expert to set up the system, and some inconsistencies. However, the users reported that they imagine that most people would learn to use this system very quickly and that various functions in the system are well integrated. These observations suggest that although the AR technology is not there yet, AR/VR has the potential to be transformative and used in human-robot collaboration tasks or similar complex tasks.

VI. Study Limitations

One limitation is that the participants recruited in this study were college students predominately seeking advanced degrees in engineering. Future work should focus on industry workers as the majority of jobs in manufacturing are taken up by high-school graduates [79]. Additionally, efforts should be made to ensure a diverse sample of older and younger in the planned participant pool. Another limitation of this study is that all the participants had no previous experience with AR. All the participants performed practice trials using the AR UI. However, the use of this new technology potentially required more mental effort to comprehend. Future studies will investigate the time-on-task effect, where the learning effect will be examined based on earlier and later trials of using the AR UI.

VII. Conclusion

In this study, we developed an AR UI to guide users in performing a pick-and-place task. The interface includes visuals that direct the user to locate an object and place the robot gripper in an optimal location for picking up an object. By conducting a user study where participants performed the pick-and-place task with the use of the AR UI and comparing them to a control group, we investigated the effects of the AR UI on task performance, cognitive workload, and situational awareness. Our findings indicate that the AR UI had a negative impact on task performance. Additionally, the AR UI significantly increased the cognitive workload perceptions and reduced situational awareness indicating that the task was more cognitively demanding when using the developed UI. However, there were no significant differences in blink rate and blink duration. These findings suggest that using the AR UI has a negative impact while performing collaborative tasks. However, the anecdotal comments from users indicated that the negative impact is primarily because of the limitations of the current AR systems and their participant's lack of familiarity with AR. We believe that advancements in AR technology have the transformative potential to be used in collaborative environments.

References

[1] F. Sherwani, M. M. Asad, and B. S. K. K. Ibrahim, "Collaborative robots and industrial revolution 4.0 (ir 4.0)," in *2020 International Conference on Emerging Trends in Smart Technologies (ICETST)*. IEEE, 2020, pp. 1–5.

[2] S. Bragança, E. Costa, I. Castellucci, and P. M. Arezes, "A brief overview of the use of collaborative robots in industry 4.0: human role and safety," *Occupational and environmental safety and health*, pp. 641–650, 2019.

[3] D. Gorecky, M. Schmitt, M. Loskyll, and D. Zühlke, "Human-machine-interaction in the industry 4.0 era," in *2014 12th IEEE international conference on industrial informatics (INDIN)*. Ieee, 2014, pp. 289–294.

[4] A. Adel, "Future of industry 5.0 in society: human-centric solutions, challenges and prospective research areas," *Journal of Cloud Computing*, vol. 11, no. 1, pp. 1–15, 2022.

[5] S. Nahavandi, "Industry 5.0—a human-centric solution," *Sustainability*, vol. 11, no. 16, p. 4371, 2019.

[6] F. Longo, A. Padovano, and S. Umbrello, "Value-oriented and ethical technology engineering in industry 5.0: A human-centric perspective for the design of the factory of the future," *Applied Sciences*, vol. 10, no. 12, p. 4182, 2020.

[7] V. Villani, F. Pini, F. Leali, and C. Secchi, "Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications," *Mechatronics*, vol. 55, pp. 248–266, 2018.

[8] R. R. Galin and R. V. Meshcheryakov, "Human-robot interaction efficiency and human-robot collaboration," in *Robotics: Industry 4.0 Issues & New Intelligent Control Paradigms*. Springer, 2020, pp. 55–63.

[9] J. G. Frederiksen, S. M. D. Sørensen, L. Konge, M. B. S. Svendsen, M. Nobel-Jørgensen, F. Bjerrum, and S. A. W. Andersen, "Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: a randomized trial," *Surgical endoscopy*, vol. 34, no. 3, pp. 1244–1252, 2020.

[10] N. M. Alzahrani, "Augmented reality: A systematic review of its benefits and challenges in e-learning contexts," *Applied Sciences*, vol. 10, no. 16, p. 5660, 2020.

[11] A. Argyrou, C. Giannoulis, A. Sardelis, P. Karagiannis, G. Michalos, and S. Makris, "A data fusion system for controlling the execution status in human-robot collaborative cells," *Procedia CIRP*, vol. 76, pp. 193–198, 2018.

[12] G. Michalos, P. Karagiannis, S. Makris, Ö. Tokçalar, and G. Chryssolouris, "Augmented reality (ar) applications for supporting human-robot interactive cooperation," *Procedia CIRP*, vol. 41, pp. 370–375, 2016.

[13] G. Michalos, N. Kousi, P. Karagiannis, C. Gkournelos, K. Dimoulas, S. Koukas, K. Mparis, A. Papavasileiou, and S. Makris, "Seamless human robot collaborative assembly–an automotive case study," *Mechatronics*, vol. 55, pp. 194–211, 2018.

[14] S. Makris, P. Karagiannis, S. Koukas, and A.-S. Matthaiakis, "Augmented reality system for operator support in human–robot collaborative assembly," *CIRP Annals*, vol. 65, no. 1, pp. 61–64, 2016.

[15] C. Gkournelos, P. Karagiannis, N. Kousi, G. Michalos, S. Koukas, and S. Makris, "Application of wearable devices for supporting operators in human-robot cooperative assembly tasks," *Procedia CIRP*, vol. 76, pp. 177–182, 2018.

[16] S. Papanastasiou, N. Kousi, P. Karagiannis, C. Gkournelos, A. Papavasileiou, K. Dimoulas, K. Baris, S. Koukas, G. Michalos, and S. Makris, "Towards seamless human robot collaboration: integrating multimodal interaction," *The International Journal of Advanced Manufacturing Technology*, vol. 105, no. 9, pp. 3881–3897, 2019.

[17] O. Danielsson, A. Syberfeldt, R. Brewster, and L. Wang, "Assessing instructions in augmented reality for human-robot collaborative assembly by using demonstrators," *Procedia CIRP*, vol. 63, pp. 89–94, 2017.

[18] J. Baumeister, S. Y. Ssin, N. A. ElSayed, J. Dorrian, D. P. Webb, J. A. Walsh, T. M. Simon, A. Irlitti, R. T. Smith, M. Kohler *et al.*, "Cognitive cost of using augmented reality displays," *IEEE transactions on visualization and computer graphics*, vol. 23, no. 11, pp. 2378–2388, 2017.

[19] T.-T. N. Do, A. K. Singh, C. A. T. Cortes, and C.-T. Lin, "Estimating the cognitive load in physical spatial navigation," in *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2020, pp. 568–575.

[20] Z. Yang, J. Shi, W. Jiang, Y. Sui, Y. Wu, S. Ma, C. Kang, and H. Li, "Influences of augmented reality assistance on performance and cognitive loads in different stages of assembly task," *Frontiers in psychology*, vol. 10, p. 1703, 2019.

[21] K. Rehring, M. Greulich, L. Bredenfeld, and F. Ahlemann, "Let's get in touch-decision making about enterprise architecture using 3d visualization in augmented reality," 2019.

[22] A. Çöltekin, I. Lochhead, M. Madden, S. Christophe, A. Devaux, C. Pettit, O. Lock, S. Shukla, L. Herman, Z. Stachoň *et al.*, "Extended reality in spatial sciences: A review of research challenges and future directions," *ISPRS International Journal of Geo-Information*, vol. 9, no. 7, p. 439, 2020.

[23] M. R. Endsley and D. J. Garland, *Situation awareness analysis and measurement*. CRC Press, 2000.

[24] R. M. Taylor, "Situational awareness rating technique (sart): The development of a tool for aircrew systems design," in *Situational awareness*. Routledge, 2017, pp. 111–128.

[25] K. Smith and P. A. Hancock, "Situation awareness is adaptive, externally directed consciousness," *Human factors*, vol. 37, no. 1, pp. 137–148, 1995.

[26] M. Naderpour, S. Nazir, and J. Lu, "The role of situation awareness in accidents of large-scale technological systems," *Process Safety and Environmental Protection*, vol. 97, pp. 13–24, 2015.

[27] M. R. Endsley and D. J. Garland, "Pilot situation awareness training in general aviation," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 44, no. 11. SAGE Publications Sage CA: Los Angeles, CA, 2000, pp. 357–360.

[28] M. R. Endsley, "Situation awareness in aviation systems," *Handbook of aviation human factors*, vol. 257, p. 276, 1999.

[29] A. Sneddon, K. Mearns, R. Flin, and R. Bryden, "Safety and situation awareness in offshore crews," in *SPE International Conference on Health, Safety, and Environment in Oil and Gas Exploration and Production*. OnePetro, 2004.

[30] M. R. Endsley, "Automation and situation awareness," in *Automation and human performance: Theory and applications*. CRC Press, 2018, pp. 163–181.

[31] D. Miller, A. Sun, and W. Ju, "Situation awareness with different levels of automation," in *2014 IEEE International Conference on systems, man, and cybernetics (SMC)*. IEEE, 2014, pp. 688–693.

[32] E. L. Hutchins, J. D. Hollan, and D. A. Norman, "Direct manipulation interfaces," *Human–computer interaction*, vol. 1, no. 4, pp. 311–338, 1985.

[33] D. Norman, *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[34] M. Baker, R. Casey, B. Keyes, and H. A. Yanco, "Improved interfaces for human-robot interaction in urban search and rescue," in *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, vol. 3. IEEE, 2004, pp. 2960–2965.

[35] J. Burke and R. Murphy, "Human-robot interaction in usar technical search: Two heads are better than one," in *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759)*. IEEE, 2004, pp. 307–312.

[36] H. K. Keskinpala, J. A. Adams, and K. Kawamura, "Pda-based human-robotic interface," in *SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme-System Security and Assurance (Cat. No. 03CH37483)*, vol. 4. IEEE, 2003, pp. 3931–3936.

[37] C. W. Nielsen and M. A. Goodrich, "Comparing the usefulness of video and map information in navigation tasks," in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, 2006, pp. 95–101.

[38] W. Boucsein, *Electrodermal activity*. Springer Science & Business Media, 2012.

[39] P. Soltani and A. H. Morice, "Augmented reality tools for sports education and training," *Computers & Education*, vol. 155, p. 103923, 2020.

[40] M. Ostanin, S. Mikhel, A. Evlampiev, V. Skvortsova, and A. Klimchik, "Human-robot interaction for robotic manipulator programming in mixed reality," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 2805–2811.

[41] S. Y. Gadre, E. Rosen, G. Chien, E. Phillips, S. Tellex, and G. Konidaris, "End-user robot programming using mixed reality," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 2707–2713.

[42] C. P. Quintero, S. Li, M. K. Pan, W. P. Chan, H. Machiel Van der Loos, and E. Croft, "Robot programming through augmented trajectories in augmented reality," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1838–1844.

[43] D. B. Kaber, J. M. Riley, K.-W. Tan, and M. R. Endsley, "On the design of adaptive automation for complex systems," *International Journal of Cognitive Ergonomics*, vol. 5, no. 1, pp. 37–57, 2001.

[44] W. C. Harris, P. A. Hancock, E. J. Arthur, and J. Caird, "Performance, workload, and fatigue changes associated with automation," *The International Journal of Aviation Psychology*, vol. 5, no. 2, pp. 169–185, 1995.

[45] R. Kumar, S. Lal, S. Kumar, and P. Chand, “Object detection and recognition for a pick and place robot,” in *Asia-Pacific world congress on computer science and engineering*. IEEE, 2014, pp. 1–7.

[46] K. Ghadge, S. More, P. Gaikwad, and S. Chillal, “Robotic arm for pick and place application,” *International Journal of Mechanical Engineering and Technology*, vol. 9, no. 1, pp. 125–133, 2018.

[47] T. Lozano-Pérez, J. L. Jones, E. Mazer, and P. A. O’Donnell, “Task-level planning of pick-and-place robot motions,” *Computer*, vol. 22, no. 3, pp. 21–29, 1989.

[48] R. Bogue Consultant, “Random bin picking: has its time finally come?” *Assembly Automation*, vol. 34, no. 3, pp. 217–221, 2014.

[49] L. F. de Souza Cardoso, F. C. M. Q. Mariano, and E. R. Zorzal, “A survey of industrial augmented reality,” *Computers & Industrial Engineering*, vol. 139, p. 106159, 2020.

[50] F. Lamberti, F. Manuri, A. Sanna, G. Paravati, P. Pezzolla, and P. Montuschi, “Challenges, opportunities, and future trends of emerging techniques for augmented reality-based maintenance,” *IEEE Transactions on Emerging Topics in Computing*, vol. 2, no. 4, pp. 411–421, 2014.

[51] D. Van Krevelen and R. Poelman, “A survey of augmented reality technologies, applications and limitations,” *International journal of virtual reality*, vol. 9, no. 2, pp. 1–20, 2010.

[52] A. Deshpande and I. Kim, “The effects of augmented reality on improving spatial problem solving for object assembly,” *Advanced Engineering Informatics*, vol. 38, pp. 760–775, 2018.

[53] S. Stadler, K. Kain, M. Giuliani, N. Mirnig, G. Stollnberger, and M. Tscheligi, “Augmented reality for industrial robot programmers: Workload analysis for task-based, augmented reality-supported robot control,” in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2016, pp. 179–184.

[54] S. Kim, M. A. Nussbaum, and J. L. Gabbard, “Influences of augmented reality head-worn display type and user interface design on performance and usability in simulated warehouse order picking,” *Applied ergonomics*, vol. 74, pp. 186–193, 2019.

[55] E. H. Hess and J. M. Polt, “Pupil size in relation to mental activity during simple problem-solving,” *Science*, vol. 143, no. 3611, pp. 1190–1192, 1964.

[56] D. Kahneman and J. Beatty, “Pupil diameter and load on memory,” *Science*, vol. 154, no. 3756, pp. 1583–1585, 1966.

[57] J. Bradshaw, “Pupil size as a measure of arousal during information processing,” *Nature*, vol. 216, no. 5114, pp. 515–516, 1967.

[58] H. Simpson, “Effects of a task-relevant response on pupil size,” *Psychophysiology*, vol. 6, no. 2, pp. 115–121, 1969.

[59] W. W. Wierwille, M. Rahimi, and J. G. Casali, “Evaluation of 16 measures of mental workload using a simulated flight task emphasizing mediational activity,” *Human factors*, vol. 27, no. 5, pp. 489–502, 1985.

[60] C. Berka, D. J. Levendowski, M. M. Cvetinovic, M. M. Petrovic, G. Davis, M. N. Lumicao, V. T. Zivkovic, M. V. Popovic, and R. Olmstead, “Real-time analysis of EEG indexes of alertness, cognition, and memory acquired with a wireless EEG headset,” *International Journal of Human-Computer Interaction*, vol. 17, no. 2, pp. 151–170, Jun. 2004.

[61] S. Chen, J. Epps, N. Ruiz, and F. Chen, “Eye activity as a measure of human mental effort in hci,” in *Proceedings of the 16th international conference on Intelligent user interfaces*, 2011, pp. 315–318.

[62] M. K. Holland and G. Tarlow, “Blinking and mental load,” *Psychological Reports*, vol. 31, no. 1, pp. 119–127, 1972.

[63] D. Kahneman, J. Beatty, and I. Pollack, “Perceptual deficit during a mental task,” *Science*, vol. 157, no. 3785, pp. 218–219, 1967.

[64] J. Beatty and B. L. Wagoner, “Pupillometric signs of brain activation vary with level of cognitive processing,” *Science*, vol. 199, no. 4334, pp. 1216–1218, 1978.

[65] J. Beatty, “Task-evoked pupillary responses, processing load, and the structure of processing resources.” *Psychological bulletin*, vol. 91, no. 2, p. 276, 1982.

[66] S. T. Iqbal, X. S. Zheng, and B. P. Bailey, “Task-evoked pupillary response to mental workload in human-computer interaction,” in *CHI’04 extended abstracts on Human factors in computing systems*, 2004, pp. 1477–1480.

[67] T. Bafna, J. P. P. Hansen, and P. Baekgaard, “Cognitive load during eye-typing,” in *ACM Symposium on Eye Tracking Research and Applications*, ser. ETRA ’20 Full Papers. New York, NY, USA: Association for Computing Machinery, 2020. [Online]. Available: https://doi.org/10.1145/3379155.3391333

[68] R. Goldstein, L. O. Bauer, and J. A. Stern, “Effect of task difficulty and interstimulus interval on blink parameters,” *International Journal of Psychophysiology*, vol. 13, no. 2, pp. 111–117, 1992.

[69] E. Galy, “Consideration of several mental workload categories: perspectives for elaboration of new ergonomic recommendations concerning shiftwork,” *Theoretical issues in ergonomics science*, vol. 19, no. 4, pp. 483–497, Sep. 2018.

[70] M. S. Young and N. A. Stanton, “Attention and automation: new perspectives on mental underload and performance,” *Theoretical issues in ergonomics science*, vol. 3, no. 2, pp. 178–194, 2002.

[71] L. M. Mazur, P. R. Mosaly, C. Moore, and L. Marks, “Association of the usability of electronic health records with cognitive workload and performance levels among physicians,” *JAMA network open*, vol. 2, no. 4, pp. e191 709–e191 709, Apr. 2019.

[72] H. Ayaz, B. Willems, B. Bunce, P. A. Shewokis, K. Izzetoglu, S. Hah, A. Deshmukh, and B. Onaral, “Cognitive workload assessment of air traffic controllers using optical brain imaging sensors,” *Advances in understanding human performance: Neuroergonomics, human factors design, and special populations*, pp. 21–31, May 2010.

[73] G. R. J. Hockey, “Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework,” *Biological psychology*, vol. 45, no. 1-3, pp. 73–93, Mar. 1997.

[74] Wickens, Christopher D and McCarley, Jason S, *Applied attention theory*. CRC press, 2019.

[75] S. Ruano, C. Cuevas, G. Gallego, and N. García, “Augmented reality tool for the situational awareness improvement of uav operators,” *Sensors*, vol. 17, no. 2, p. 297, 2017.

[76] C. Reardon, J. Gregory, C. Nieto-Granda, and J. G. Rogers, “Enabling situational awareness via augmented reality of autonomous robot-based environmental change detection,” in *International Conference on Human-Computer Interaction*. Springer, 2020, pp. 611–628.

[77] D. Aschenbrenner, N. Maltry, J. Kimmel, M. Albert, J. Scharnagl, and K. Schilling, “Artab-using virtual and augmented reality methods for an improved situation awareness for telemaintenance,” *IFAC-PapersOnLine*, vol. 49, no. 30, pp. 204–209, 2016.

[78] M. Wallmyr, T. A. Sitompul, T. Holstein, and R. Lindell, “Evaluating mixed reality notifications to support excavator operator awareness,” in *IFIP Conference on Human-Computer Interaction*. Springer, 2019, pp. 743–762.

[79] US Census Bureau, “Occupational employment projections to 2022,” https://www.bls.gov/opub/mlr/2013/article/pdf/occupational-employment-projections-to-2022.pdf, Dec. 2013.