Using Augmented Reality in Human-Robot Assembly: A Comparative Study of Eye-Gaze and Hand-Ray Pointing Methods

Sławomir K. Tadeja[1], Tianye Zhou[1], Matteo Capponi[1,2], Krzysztof Walas[3,4], Thomas Bohné[1], and Fulvio Forni[1]

***Abstract*— Collaborative robots (cobots) are a promising technology for frontline workers in industry. They can support tasks that cannot be fully automated but are repetitive, fatiguing, boring, or dangerous for humans. Although cobots are explicitly designed to work with humans, they remain primarily non-intuitive and difficult to collaborate with. Thus, there is a need for new interaction approaches to facilitate efficient human-robot collaboration. Recently, we could see emerging examples of using *augmented reality* (AR) to assist a worker in collaborative task execution with a cobot. However, for such an approach to provide truly efficient support for the seamless bimanual task execution, we need to first investigate interaction methods offered by an AR interface. To that end, we performed a study with sixteen participants to compare eye-gaze and hand-ray pointing methods for part selection in collaborative, manual assembly tasks. The results of our study show that both techniques provide similar perceived usability, with the eye-gaze selection leading to significantly shorter completion times.**

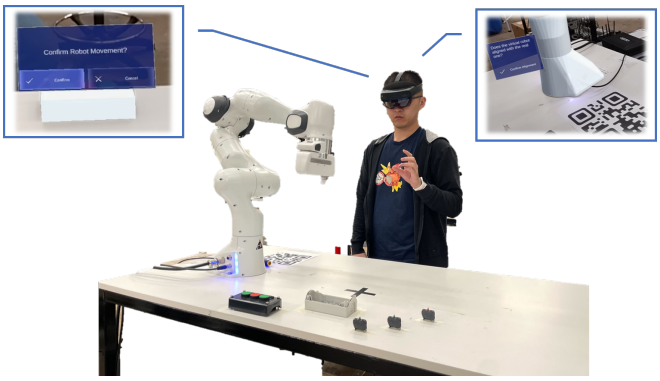

Fig. 1. A user is collaborating with the robotic arm (i.e., a cobot) on the manual assembly of a control station using a see-through AR headset.

I. INTRODUCTION

Collaborative robots (cobots) are often tasked to work alongside humans in shared workspaces. Their design emphasizes safety (bounded motion and forces, possibly reduced inertia and back-drivability). They can improve tasks in manufacturing that cannot be fully automated but are repetitive, fatiguing, tiresome, or dangerous for humans, who instead could focus on more complex or higher value-adding work [1]. If done successfully, cobots combined with the abilities of humans can increase production quality, safety, and well-being while at the same time reducing waste and production costs [1].

However, the integration of cobots onto manufacturing shop floors is currently held back by poor human-robot collaboration (HRC) [2]. Instead of helpful technology designed to improve working conditions, people on the shop floor often perceive robots as suspect, difficult, and non-intuitive to work with [3]. To address these issues, it has been proposed to use augmented reality (AR) [1], [4], [5], [6] as a promising approach to facilitating more intuitive HRC [1], [4].

While using AR for HRC seems promising, research on this topic is scarce [5], [6], [7]. Visual information exchange with AR in robotics is still in its infancy. There are only very few examples where AR-based human-robot interaction (HRI) has been attempted. These include preliminary studies and proofs-of-concept in which the AR experience focuses on hand-held [5], or head-worn [8], [9], [10] devices to provide a suite of functionalities typically associated with teach pendants for remote robot operation [11]. Although these initial attempts are valuable, they have not directly addressed the problem of designing an effective and plausible human interaction method when using AR to engage in worker-cobot joint task execution.

To advance AR-based HRI, we propose a novel system for bi-directional information exchange between a human and a robot [4]. In such context, AR can be used to, for instance, visually exchange information to increase trust during tasks requiring human-robot teaming [12]. In contrast to previous research, we performed extensive user studies to ground our findings. In our study, we focused on a collaborative manual assembly scenario in which a human and cobot are working side by side, specifically where the cobot is responsible for picking & placing parts and components of an industrial control station that then has to be manually assembled by a worker. To facilitate HRI, we contrasted two AR-based techniques for parts selection utilizing (A) eye-gaze and (B) hand-ray pointing, respectively. In what follows below, we present the results of two studies: a formative, observational study with domain experts used to ground our work and provide an initial assessment of both modalities. A quantitative, repeated-measure user study with sixteen participants was conducted to reason and compare both HRI techniques. The study results suggest a statistically significant shortened time of assembly part selection with

*This work was supported by a grant from the *Observatory for Human Machine Collaboration (OHMC)* at the University of Cambridge, UK. The research of S.K. Tadeja and T. Bohné was supported by EPSRC grant no. EP/V062123/1

[1]University of Cambridge, [2]Politecnico di Torino, [3]Poznań University of Technology, [4]IDEAS NCBR. Contact: skt40@eng.cam.ac.uk

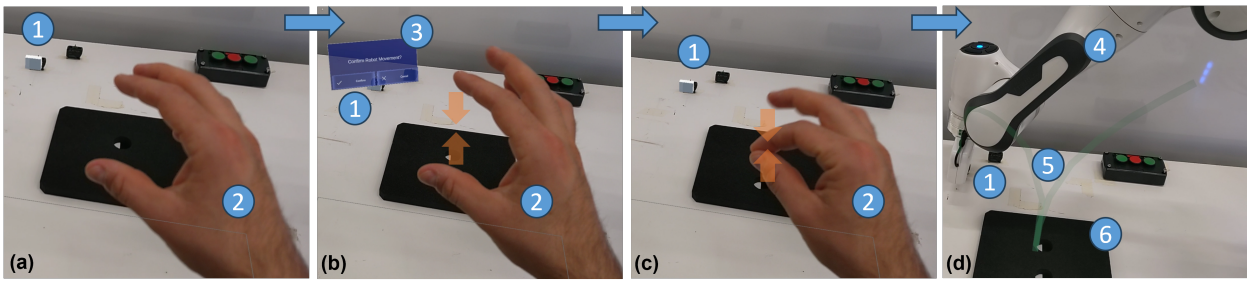

Fig. 2. (a-d) The user collaboratively assembles a control station using eye-gaze pointing. (a) The operator gazes over the desired (a-d)(1) part, which recalls a confirmation (b)(3) pop-up. Next, the operator gazes over [yes] button and executes (b)(2)-(c)(2) the *pinch gesture*. This leads to (d)(4) cobot grabbing the (d)(1) part, which is transported over (d)(5) the visualized movement trajectory to (d)(6) the placement zone.

(A) eye-gaze approach. The study also showed that in a relatively short time, human operators exhibited a change in their behavior and were comfortable engaging with their task before the cobot completely finished its movement. We could also see that 25% of all AR operators started to parallelize their task execution when collaborating with the robot. In summary, our contributions are threefold:

1) The AR-based system supports joint human-cobot task execution through bi-directional information exchange of the task intentions (e.g., human part selection and hand tracking, and robot movement trajectory).
2) Increased operator safety achieved through collision avoidance utilizing AR-based hand tracking.
3) In-depth evaluation of eye-gaze and hand-ray pointing interaction during HRC task based on a formative study with domain experts and subsequent extensive user experiment.

II. RELATED WORK

Integrating AR technology with robotics remains an emerging and underexplored field with limited research studies published to date [4], [7]. However, despite a few examples, the growing body of relevant existing literature highlights the potential of AR to facilitate HRI [4], [6], [13].

Groechel et al. [14] recognized the need for research in immersion-based HRI and developed a comprehensive tool to organize key features of such systems, providing a consistent classification method and insights into emerging trends and future works. Moreover, the authors underline the need for user experiments to gain new insights about HRI facilitated with immersive technologies.

Hietanen et al. [15] built an interactive AR interface to ensure safety in collaborative manufacturing processes with cobots. Also, Szczurek et al. [16] introduced a novel human-robot interface combining 3D mixed reality with operator physiological parameters monitoring to enhance safety and task efficiency. In other instances, Szczurek et al. [11] also developed a mixed reality remote control system to ensure high-precision interaction and specialized visualization suitable for particle accelerator facilities.

Some researchers have also investigated using AR when programming robots [8], [9], [17], [18], [19]. For example, Kapinus et al. [18] developed an innovative end-user interface for programming space robotic tasks in industrial contexts. Moreover, Thoo et al.[17] proposed an immersive, virtual representation of the real workspace to facilitate robot programming without needing direct proximity to the actual machine. Also, Blankemeyer et al. [9] presented a new AR-based method for intuitive robot programming, enabling robots to replicate recorded virtual assembly processes efficiently. Other authors have also studied robot programming with the help of AR in the context of improving the flexibility of robotic programming [8], using mobile AR to manage robotic workstations [19] or efficient task planning and allocation [20].

Chu et al. [21] conducted an experiment using AR to show robot intent and motion in an assembly task with LEGO blocks where the cobot and operator worked independently from each other. The AR was used to present either the robot's working envelope or the near-future position of the gripper and its movement. The study showed the potential of AR for HRC. Nonetheless, this work differs substantially from the goal of our research, where we compared two AR-based interaction techniques and the cobot and operator worked collaboratively on assembling the same object.

Although these studies offer initial results and explorations in using AR to facilitate HRI when executing collaborative tasks [15], more in-depth, interdisciplinary user studies investigating individual aspects of HRC are needed [14]. As previously advocated, the interaction method selection is crucial for any AR system to offer a compelling and effective user experience and consequently contribute to the wider adoption of AR [22]. To that end, we extend previous research by directly comparing varying interaction modalities frequently used in AR-based user studies, i.e., eye-gaze and hand-ray pointing [23]. In agreement with other researchers [14], we believe that such in-depth studies, providing both subjective (e.g. questionnaires and user feedback) and objective results (e.g., body tracking data or captured task completion time), are needed to provide foundations for adequate and meaningful AR-supported HRI.

III. HUMAN-ROBOT INTERACTION SYSTEM

We propose a novel system which couples AR and robotics to facilitate bi-directional information exchange between a worker and a cobot [4]. Our system relies on the two domi-

nant interaction methods for AR headsets, i.e., (A) eye-gaze and (B) hand-ray pointing [23], rarely used in such a context. Although these two methods are extensively researched in the human-computer interaction (HCI) field [23], there are only a few examples of where these techniques were explored in AR coupled with robotics. Here, we investigated how these two techniques could be used for part selection, after which the robotic arm moves over a predefined trajectory to grab and move the selected part to the placement zone (see Fig. 2(d)). We, therefore, compared eye-gaze with hand-ray pointing to facilitate AR-based HRC in a real-life industrial scenario:

(A) Eye-gaze pointing–When the user gazes over an assembly part, the object highlights. Next, after executing the pinch gesture, a text message pops up requesting the user to confirm the selection with the pinch gesture (see Fig. 2).

(B) Hand-rays pointing–Instead of eye-gazing, the user could also point mid-air towards the assembly part, which would cause it to be highlighted. Then, after executing the pinch gesture, a text message pops up requesting the user to confirm the selection with another pinch gesture (see Fig. 4).

A. Apparatus

We designed our system around the state-of-the-art *Microsoft HoloLens 2* (HL2) head-mounted display [24]. This see-through headset offers built-in eye-gaze supported by 2 IR sensors and hand-tracking through 4 visible light cameras.

The cobot used in our study was the *Franka Emika Research* arm robot with up to 3 kg payload. The robot has 7 degrees of freedom and was equipped with *Franka Hand* gripping end effector.

Our software was built around *Microsoft Mixed Reality Toolkit* [24] and *Unity* game engine [25], which are frequently used platforms for AR development [24], [25]. The HL2 device and the Franka Emika robot communicated over a local WiFi network using TCP/IP protocol and Robot Operating System (ROS).

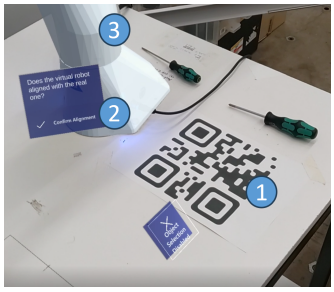

Fig. 3. (1) A QR code and (2) textual instructions were used to calibrate and (3) align the cobot with its model to ensure correct robot tracking [5].

B. System Calibration

To provide a bi-directional information exchange between the cobot and the AR headset, the system must be initially calibrated to allow both devices to work within the same reference frame. We achieve this using a QR code that can be natively detected and read by HL2 [24] (see Fig. 3). Such an approach provided a robust correct positioning and tracking method within the same coordinate system [5], [13].

C. Collision Avoidance with AR

After calibrating the system, the cobot's digital model is rendered invisible and precisely aligned with its physical counterpart within the HL2 coordinate system. We also placed virtual box colliders near the worker's head and hands as tracked by the headset. This native-to-game engine mechanism serves as a proactive measure to detect and prevent potential contact between the headset wearer and the cobot. Upon detecting a potential collision, the system immediately halts robot operations, resuming only when the collision threat is eliminated, thereby ensuring operator safety without compromising the immersive AR experience. Such an approach contributes further towards the efforts of providing a truly cooperative environment by preventing any direct human-robot contact.

D. Robot Trajectory Generation

The movement trajectory is generated using B-splines, i.e., a set of piecewise polynomial functions [26], utilizing control points derived from the positions of the robot's end effector, the center of assembly part, and the middle of the placement zone. The B-spline's adherence to the convex hull property [27] ensures the trajectory remains within the bounds set by the control points, effectively mitigating risks of intruding into worker's operational areas. Moreover, the trajectory was rendered as semi-transparent to avoid information overload [28] (see Fig. 2(d) and Fig. 4(d)).

E. Robot Controller

The robot employs a Cartesian impedance controller [29] leveraging PID (Proportional, Integral, Derivative) control to emulate the effects of springs and dampers plus the bounded adaptive action of the integral component. The controller connects these virtual mechanisms from the robot's end effector to a three-dimensional target position, ensuring reaching the intended location. Such a setup allows the robot to adapt to external forces by simulating mechanical impedance. Therefore, when a participant accidentally or purposely pushes the cobot from its path, the controller ensures it can autonomously return to and continue along its predefined trajectory.

IV. FORMATIVE USER STUDY

We decided to conduct a formative study with domain experts [30] to pinpoint and fix issues with our experimental and HRI designs before commencing large-scale experiment.

A. Participants

We invited six domain-expert participants, understood as having prior experience with either robotics and control, AR/VR or both hereinafter referred to as **P1–P6**, to take part in our usability study [30]:

P1: was a 23-year-old PhD student in control engineering. He holds an undergraduate degree in aerospace engineering and a postgraduate qualification in control engineering. He had never used AR/VR devices but had relevant experience working with robots.

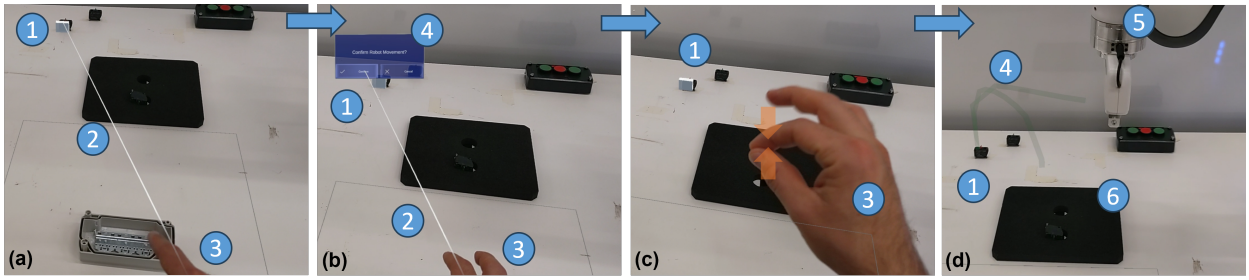

Fig. 4. (a-d) The user collaboratively assembles a control station using hand-ray pointing. (a) The operator points towards the desired (a-d)(1) part, which recalls the confirmation (b)(4) pop-up. Next, the operator points with (b)(2) ray towards the [yes] button and executes (c)(3) the *pinch gesture*. This leads to (d)(5) cobot moving to grab the (d)(1) part, which is transported over (d)(4) the visualized movement trajectory to (d)(6) placement zone.

P2: was a 25-year-old PhD student in control engineering. His undergraduate and graduate degrees were in information engineering. He had never used AR but reported having four years of experience with VR gaming and having also been working with cobots for the same time.

P3: was a 44-year-old professor of information engineering and neuroscience. He had never worked with robots before nor with AR and VR technologies.

P4: was a 26-year-old PhD student in engineering. He works on control systems and holds undergraduate and graduate degrees in information engineering. He had never worked with cobots and had limited exposure to VR.

P5: was a 27-year-old PhD student in neuroscience. She reported having no prior experience working with cobots. She also admitted to using AR once in the past.

P6: was a 41-year-old assistant professor of robotics and machine intelligence with 4-year experience in industrial manufacturing. He had relevant experience with cobots and VR/AR.

B. Tasks

The participants were tasked with collaboratively assembling a five-part control station with the help of a cobot by selecting each component using prepared HRI methods, i.e., (A) eye-gaze or (B) hand-ray pointing (see Fig. 2 and Fig. 4). The cobot performed all pick & place operations while the AR user carried out the assembly activities.

C. Study Protocol

The formative user study consisted of performing a collaborative assembly of an industrial asset using either (A) eye-gaze or (B) hand-ray pointing. The experiment started with an introduction that explained the aim and assembly task. Here, we collected information about participants' backgrounds and previous experience with cobots and immersive technologies. Next, the participants underwent training sessions to learn our two interaction methods, ensuring that the context differed from HRC to avoid premature familiarity.

Immediately after the training, participants performed two assembly trials in alternating selection modalities. The participants were also informed to either wait for the robot to halt before they performed manual assembly or carry out activities in parallel with the robot's movements.

Each trial ended with participants filling in the *NASA Task Cognition Load* (TLX) [31], and *System Usability Scale* (SUS) [32]. The latter is often employed to assess the perceived usability of a given interface and was used to express participants' satisfaction and difficulty. Meanwhile, the TLX captured subjective cognitive load when performing a given task using an immersive system [33], [34]. We concluded the experiment by conducting another interview to gather unstructured feedback and potential improvements. Furthermore, all the participants were asked to fill in the *Simulation Sickness Questionnaire* (SSQ) [35] before and after the experiment. This questionnaire is commonly administered to understand if a given immersive system induces simulation sickness symptoms [33], [34].

In terms of objective data, we collected a series of time measurements reported in Tab. II and Tab. III, including (1) **Completion Time** representing the total duration of the assembly process; (2) **Interaction Time** denoting the time in which the participant and the cobot perform activities together between the part selection starts and when the robot finishes its movement; and (3) **Total Selection Time**, understood as the duration the participants took to select the various components, defined as the sum of the time intervals between the start of the part selection process until the robot movement confirmation. Such an approach helped us focus on the comparison of HRI methods and decouple the results from the unassisted control station manual assembly task and cobot movement interruptions.

V. FORMATIVE STUDY RESULTS

A. Questionnaires

In the pilot study with six participants, the initial results indicate a balance in perceived workload between conditions (A) and (B), with half of the participants finding the former more demanding and vice-versa. Moreover, P1 reported a much higher workload for (A) $(81.7 > 54.3)$ while P3 perceived a similar, high workload in both conditions (see Table I). When interpreting the individual NASA TLX results [36], we can consider P1's and P3's results as "very high" (> 80) and "high" $(50 - 79)$ for (A) and (B), respectively. This may suggest their perceived workload as an outlier, especially since P3 also reported other issues. The remaining scores are mixed of either "medium" $(10 - 29)$ (P4, P5),

"somewhat high" $(30-49)$ (P5, P7) and "high" (>50) (P5) for (A) and (B).

Satisfaction ratings varied, with four participants, i.e., P3-P6, indicating a lower satisfaction for condition (B), with P3 being dissatisfied with both selection methods and rating them as fairly difficult. In general, both selection techniques were rated similarly, with no trend of one over the other.

The SSQ showed no potential for simulation sickness symptoms being induced by our HRI methods. The average SSQ score before the experiment was 3.74 and after the experiment was 13.71. Both values were below the generally accepted level of concern of 20.0 [37], indicating the manageable impact of our AR interface on users.

TABLE I
THE RESULTS OF NASA TLX AND SFS QUESTIONNAIRES.

	Age [years]	**Condition** (A or B)	**TLX** [0-100] ↓	**Satisfaction** [1-7] ↑	**Difficulty** [1-7] ↓
P1	23	A	81.7	3	6
		B	54.3	5	5
P2	25	A	30.0	6	2
		B	26,7	6	2
P3	44	A	75.0	3	5
		B	77.3	2	6
P4	26	A	11.3	7	2
		B	21.3	6	2
P5	27	A	35.3	4	5
		B	52.3	3	4
P6	41	A	54.7	3	5
		B	46.3	1	5

B. Observed Behavior

The formative study with domain experts helped us to pinpoint and resolve some design faults. First, participants P3, P5, and P6 had recurring problems correctly performing the pinch gesture that HL2 did not always detect.

Secondly, none of the participants paralleled their activities with those of the robot, i.e., they always waited for the robot to stop moving before grabbing and assembling manually the new component. This could be caused by participants having insufficient exposure to AR and the collaborative assembly conditions where the robotic arm was moving near a person.

Finally, an important finding was the challenge some participants faced during the manual assembly activities. The HL2 used finger markers to continuously provide feedback about hand movement tracking. However, participants P3, P4, and P5 reported these obstructed their field of view, leading to increased fatigue. The obstruction was particularly problematic when placing the three small components on the base and when attaching the cover of the control station.

C. Suggested Changes

The study results helped us to refine the interface for a subsequent large-scale experiment. We refrained from statistical analysis of the gathered data, as the results could be misleading with such a small sample size. Instead, we focused on more qualitative observations concerning user behavior and expert feedback, often practiced in such research [30], [34]. Consequently, we optimized our interface and experimental procedure by following refinements:

1) We integrated a highlight effect for the buttons users gazed at to increase user awareness during HRI.
2) We decided to remove the visual markers of tracked fingers and visual cues about forbidden zones for the pinch gesture. They were both considered distracting and overwhelming for the experts.
3) To help users navigate the interface, a reminder saying "object selection disable" was introduced when users gazed at parts without enabling the object selection first.
4) Finally, we emphasized training the pinch gesture to ensure seamless user interaction in subsequent trials.

TABLE II
FORMATIVE STUDY: MEAN AND STD DEVIATION OF COMPLETION, INTERACTION AND TOTAL SELECTION TIMES DATA BY TRIAL TYPE

Trial type	**Size**	**Variable**	**Mean [s]**	**SD [s]**
A	6	Completion Time	298.1	58.2
B	6	Completion Time	364.6	193.8
A	6	Interaction Time	224.3	54.2
B	6	Interaction Time	325.6	204.9
A	6	Total Selection Time	68.2	38.5
B	6	Total Selection Time	105.8	84.8

VI. QUANTITATIVE USER STUDY

After refining our AR system based on formative study results, we conducted another experiment following a similar protocol with the same task and a new set of participants.

A. Participants

For this study, we relied on opportunity sampling when recruiting sixteen participants. Ten of them were male, while six were female. They were between 20-40 years old. We chose this approach as we were interested in quantitative A/B testing results to reason about the efficiency and generalizability of our selection methods rather than expert feedback.

B. Protocol

The A/B study protocol was amended based on the feedback and analysis of the domain expert study. Firstly, the duration of training for the pinch gesture was increased to address problems experienced by domain experts. The number of trials was doubled from two to four to mitigate the potential novelty effects and participants' unfamiliarity with the AR and cobots. The order of trials was fully balanced.

In addition, the *Short Flow Scale* (SFS) [38] questionnaire was administered after each trial to gain a deeper view of the participants' experiences to reflect on engagement and skill in performing the assembly task [39], [34].

VII. QUANTITATIVE USER STUDY RESULTS

A. Task Completion Times

The task completion times for the quantitative study can be found in Tab. III. Compared to the formative study results, they indicated decreased values across all measurements.

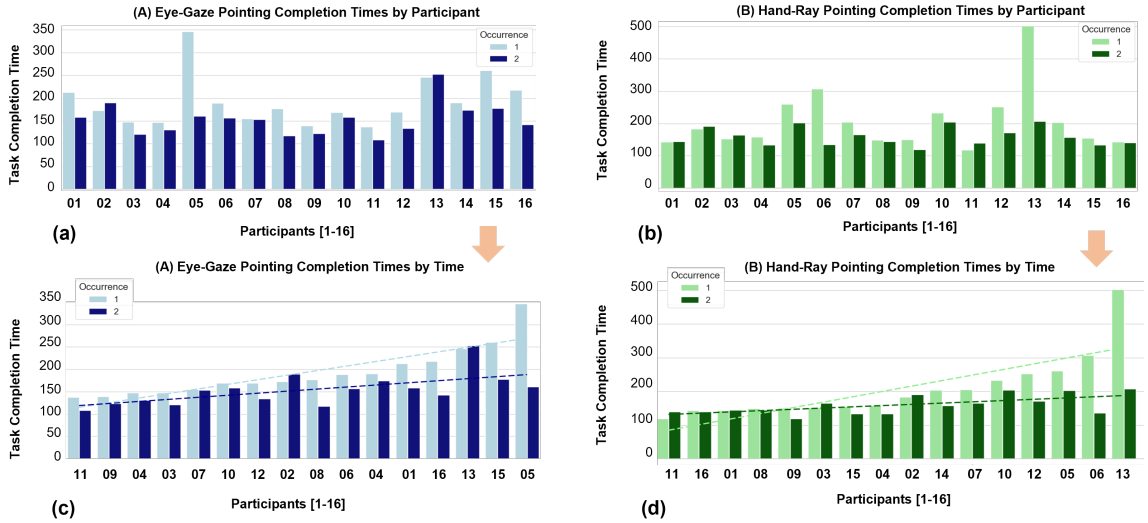

Fig. 5. (a) and (b) shows task completion times for the first and second occurrence of each trial type. Meanwhile, (c) and (d) present the same data in sorted order with fitted lines, showing that the second trial provided more homogeneous results.

Thus indicating the effectiveness of the interface improvements made based on the domain expert study results.

TABLE III

QUANTITATIVE STUDY: MEAN AND STD DEVIATION OF COMPLETION, INTERACTION AND TOTAL SELECTION TIMES DATA BY TRIAL TYPE

Trial type	Size	Variable	Mean [s]	SD [s]
A	32	Completion Time	172.8	49.6
B	32	Completion Time	183.3	72.8
A	32	Interaction Time	142.6	45.2
B	32	Interaction Time	155.4	64.5
A	32	Total Selection Time	39.0	20.7
B	32	Total Selection Time	47.5	33.4

Fig. 5(a)(b) and show the participants' completion times, separated by interaction mode (A) eye-gaze or (B) hand-ray pointing, respectively. The results indicate learning between the first and the second trial in both cases, as completion time decreased in 75% (12/16) of participants for both modalities. We can also see that the completion times of the second occurrence of a given trial type tend to be lower and more homogeneous among participants. To illustrate this, Fig. 5(c)(d) shows the ordered task completion times with fitted lines indicating more stable results in subsequent trials of the same type across all participants.

We can attribute this behavior to improved proficiency in manual assembly tasks and decreased total selection time. However, as these data are influenced by the manual skills of each participant, the following analysis focused only on selection time to assess the effectiveness of each interaction modality. We decided on such an approach to isolate the impact of each participant's manual skill learning, ensuring a more accurate comparison of the two interaction modes.

We analyzed the data using the Wilcoxon post hoc signed-rank-sum test as the data was non-parametric, and the experiment followed a repeated-measures design with two

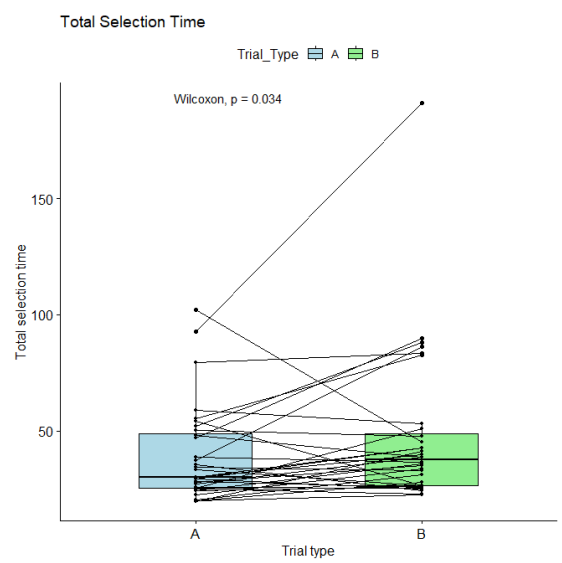

Fig. 6. Paired boxplot of total selection times by trial type.

conditions [40] (see Fig. 6). The test showed the statistically significant result of $W(32) = 151, p = 0.034$ for the total selection time, providing evidence that there is an observable difference between the two modalities in the crucial objective measurement directly comparing eye-gaze and hand-ray selection with the results favoring the former.

B. Questionnaires

In terms of mean values, the average workload score for (A) was equal to 26.7, while for (B), it was 28, suggesting a slightly higher perceived workload for the hand-ray pointing. Further granularity is provided by the box plot in Fig. 7, which shows that scores for (A) clustered more closely together, indicating a more uniform user experience. In contrast, (B) shows a wider spread of scores, reflecting a

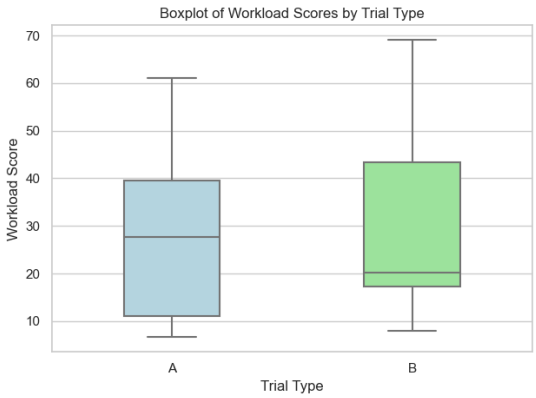

Fig. 7. Boxplot of TLX results by trial type.

heterogeneous user response. This variability in (B) could mean that while some users may find the hand-ray pointing intuitive and relatively effortless, others may find it challenging, resulting in a wider range of workload scores.

Conversely, despite its slightly higher average workload score, the eye-gaze pointing appeared to provide a more consistent user experience. This result was also confirmed by the shorter selection times, highlighting a smoother experience compared to the hand-ray pointing approach.

The SUS and SFS questionnaires revealed no significant differences between the two modalities, suggesting that both exhibited comparable usability, flow, and anxiety states. Moreover, the SFS scores for trial types had the same median of 5.3 and 3.5 for flow and anxiety levels, respectively. Regarding usability, it is worth noting that both for (A) and (B) modalities, the average SUS score exceeded the minimum 70% threshold (i.e., $SUS_A = 79.3\%, SUS_B = 80.2\%$), respectively, traditionally recognized as a benchmark for good usability and satisfactory user experience [32].

Furthermore, we also asked participants to fill in the SSQ before and after the whole experiment to assess potential health issues it may cause. The average total SSQ score was 4.2 before the experiment. After using the AR interface, the score increased to 12.4, still well below the aforementioned critical threshold of 20.0 [37].

C. Observed Behavior

Unlike the formative study, most participants seemed comfortable performing tasks while the robot was still completing its movement. From the second trial onward, they tended to pick up the component from the placement zone before the robot returned to its initial position. Moreover, four people asked the robot for the next component while still engaged in the manual assembly activity of the previous component, thus paralleling the robot's pick & place operation with the operator's manual assembly task. This tentatively indicates confidence and comfort in working with robotic systems, suggesting that, over time, individuals may develop work patterns that are closely synchronized with robotic operations. Only two participants continuously waited for the robot to stop moving. Interestingly, one of them had consistent prior robotic experience and decided to keep away to avoid collisions or time loss due to an emergency stop.

Two participants experienced significant difficulties with the (B) hand-ray pointing modality, finding it difficult to keep the hand-ray within their respective fields of view. As a result, they tended to step back from the work area to gain visual space. In contrast to the pilot study, where half of the participants reported difficulties in performing the pinch gesture, in the subsequent study, only one participant had serious difficulties with performing this gesture.

VIII. CONCLUSIONS

Combining AR with robotics can provide intuitive and real-time information-sharing capabilities that may increase workers' safety and efficiency of joint task execution [1]. Moreover, AR can provide an interface for humans and cobots that simplifies information exchange and fast collaboration in complex industrial environments (see Fig. 1). The participants found the AR interface useful and expressed a sense of safety while interacting with the robot. This increases the promise of integrating AR technology into future HRC systems.

However, as in the case of any new interface and use case scenario, coupling head-worn AR with robots requires extensive user studies to help designers select and develop the most efficient approaches to HRI. To that end, our paper provides one of the first in-depth insights into two common AR selection methods based on eye-gaze (see Fig. 2) and hand-ray pointing (see Fig. 4) used within HRC task. The result of the user study suggests that using eye-gaze for object selection in collaborative assembly task is not only favored by the users but also leads to statistically significantly faster part selection times (see Fig. 6).

In the future, we plan to study in-depth the eye-gaze-based interaction and its influence on the HRC. Specifically, we plan to compare the behavior of participants in non-AR and AR-supported scenarios to investigate the effects of visualizing robot trajectory as well as the safety provided by the collider mechanism. We also plan to explore whether coupling with other communication methods, such as audio or environmental [4], could offer additional advantages.

REFERENCES

[1] R. Raffik, S. R. Ramamoorthy, V. Sushmita, N. J. Lakshmi, and S. Gautham, "Enhancing human cobot interaction with mixed reality: A futuristic review," in *2023 2nd ICAECA*, 2023, pp. 1–6.

[2] E. Matheson, R. Minto, E. G. G. Zampieri, M. Faccio, and G. Rosati, "Human–robot collaboration in manufacturing applications: A review," *Robotics*, vol. 8, no. 4, 2019.

[3] D. Wurhofer, T. Meneweger, V. Fuchsberger, and M. Tscheligi, "Deploying Robots in a Production Environment: A Study on Temporal Transitions of Workers' Experiences," in *Human-Computer Interaction – INTERACT 2015*, ser. Lecture Notes in Computer Science, J. Abascal, S. Barbosa, M. Fetter, T. Gross, P. Palanque, and M. Winckler, Eds. Cham: Springer International Publishing, 2015, pp. 203–220.

[4] S. A. Green, M. Billinghurst, X. Chen, and J. G. Chase, "Human-Robot Collaboration: A Literature Review and Augmented Reality Approach in Design," *International Journal of Advanced Robotic Systems*, vol. 5, no. 1, p. 1, 2008.

[5] K. Zieliński, K. Walas, J. Heredia, and M. Kjærgaard, "A Study of Cobot Practitioners Needs for AR Interfaces in the Context of Current Technologies," in *30th IEEE RO-MAN*, 2021, pp. 292–298.

[6] R. K. Khamaisi, E. Prati, M. Peruzzini, R. Raffaeli, and M. Pellicciari, "UX in AR-Supported Industrial Human–Robot Collaborative Tasks: A Systematic Review," *Appl. Sci.*, vol. 11, no. 21, p. 10448, Jan. 2021.

[7] M. Walker, T. Phung, T. Chakraborti, T. Williams, and D. Szafir, "Virtual, Augmented, and Mixed Reality for Human-robot Interaction: A Survey and Virtual Design Element Taxonomy," *ACM Transactions on Human-Robot Interaction*, vol. 12, no. 4, pp. 43:1–43:39, July 2023.

[8] 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.

[9] S. Blankemeyer, R. Wiemann, L. Posniak, C. Pregizer, and A. Raatz, "Intuitive robot programming using augmented reality," *Procedia CIRP*, vol. 76, pp. 155–160, 2018, 7th CIRP Conference on Assembly Technologies and Systems (CATS 2018).

[10] P. Gao and H. Zhang, "Bayesian Deep Graph Matching for Correspondence Identification in Collaborative Perception," in *Robotics: Science and Systems XVII*. Robotics: Science and Systems Foundation, July 2021. [Online]. Available: http://www.roboticsproceedings.org/rss17/p022.pdf

[11] K. A. Szczurek, R. M. Prades, E. Matheson, J. Rodriguez-Nogueira, and M. D. Castro, "Multimodal Multi-User Mixed Reality Human–Robot Interface for Remote Operations in Hazardous Environments," *IEEE Access*, vol. 11, pp. 17 305–17 333, 2023.

[12] M. B. Luebbers, A. Tabrez, and B. Hayes, "Augmented Reality-Based Explainable AI Strategies for Establishing Appropriate Reliance and Trust in Human-Robot Teaming," in *5th International Workshop on Virtual, Augmented, and Mixed Reality for HRI*, 2022. [Online]. Available: https://openreview.net/forum?id=SnVew9n9Tkq

[13] X. Liu, L. Zheng, Y. Wang, W. Yang, Z. Jiang, B. Wang, F. Tao, and Y. Li, "Human-centric collaborative assembly system for large-scale space deployable mechanism driven by digital twins and wearable ar devices," *J. Manuf. Syst.*, vol. 65, pp. 720–742, 2022.

[14] T. R. Groechel, M. E. Walker, C. T. Chang, E. Rosen, and J. Z. Forde, "A Tool for Organizing Key Characteristics of Virtual, Augmented, and Mixed Reality for Human–Robot Interaction Systems: Synthesizing VAM-HRI Trends and Takeaways," *IEEE Robotics & Automation Magazine*, vol. 29, no. 1, pp. 35–44, 2022.

[15] A. Hietanen, R. Pieters, M. Lanz, J. Latokartano, and J.-K. Kämäräinen, "AR-based interaction for human-robot collaborative manufacturing," *Robot. Comput.-Integr. Manuf.*, vol. 63, p. 101891, June 2020.

[16] K. A. Szczurek, R. Cittadini, R. M. Prades, E. Matheson, and M. Di Castro, "Enhanced Human–Robot Interface With Operator Physiological Parameters Monitoring and 3D Mixed Reality," *IEEE Access*, vol. 11, pp. 39 555–39 576, 2023.

[17] Y. Thoo, J. Maceiras, P. Abbet, M. Racca, H. Girgin, and S. Calinon, "Online and Offline Robot Programming via Augmented Reality Workspaces," *ArXiv*, July 2021.

[18] M. Kapinus, V. Beran, Z. Materna, and D. Bambušek, "Spatially Situated End-User Robot Programming in Augmented Reality," in *2019 28th IEEE International Conference on Robot and Human Interactive Communication*, Oct. 2019, pp. 1–8, iSSN: 1944-9437.

[19] M. Kapinus, Z. Materna, D. Bambušek, V. Beran, and P. Smrž, "ARCOR2: Framework for Collaborative End-User Management of Industrial Robotic Workplaces using Augmented Reality," 2023.

[20] M. Lippi and A. Marino, "A mixed-integer linear programming formulation for human multi-robot task allocation," in *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*, 2021, pp. 1017–1023.

[21] C.-H. Chu and Y.-L. Liu, "Augmented reality user interface design and experimental evaluation for human-robot collaborative assembly," *Journal of Manufacturing Systems*, vol. 68, pp. 313–324, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0278612523000717

[22] B. Hari Chandana, N. Shaik, and P. Chitralingappa, "Exploring the frontiers of user experience design: Vr, ar, and the future of interaction," in *2023 International Conference on Computer Science and Emerging Technologies (CSET)*, 2023, pp. 1–6.

[23] M. N. Lystbæk, P. Rosenberg, K. Pfeuffer, J. E. Grønbæk, and H. Gellersen, "Gaze-Hand Alignment: Combining Eye Gaze and Mid-Air Pointing for Interacting with Menus in Augmented Reality," *Proc. ACM Hum.-Comput. Interact.*, vol. 6, no. ETRA, may 2022.

[24] S. K. Tadeja, D. Janik, P. Stachura, M. Tomecki, and K. Walas, "Design of ARQ: An Augmented Reality System for Assembly Training Enhanced with QR-Tagging and 3D Engineering Asset Model," in *2022 IEEE VRW*, 2022, pp. 466–471.

[25] E. Coronado, S. Itadera, and I. G. Ramirez-Alpizar, "Integrating virtual, mixed, and augmented reality to human & robot interaction applications using game engines: A brief review of accessible software tools and frameworks," *Applied Sciences*, vol. 13, no. 3, 2023.

[26] W. Böhm, "Efficient evaluation of splines," *Computing*, vol. 33, no. 2, pp. 171–177, June 1984.

[27] A. Chaudhuri, "B-Splines," in *Encyclopedia of Computer Graphics and Games*, N. Lee, Ed. Cham: Springer International Publishing, 2024, pp. 242–252.

[28] K. D. Chandan, J. Albertson, and S. Zhang, "Learning visualization policies of augmented reality for human-robot collaboration," in *Proceedings of The 6th Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, K. Liu, D. Kulic, and J. Ichnowski, Eds., vol. 205. PMLR, 14–18 Dec 2023, pp. 1233–1243. [Online]. Available: https://proceedings.mlr.press/v205/chandan23a.html

[29] A. Albu-Schaffer, C. Ott, and G. Hirzinger, "A passivity based Cartesian impedance controller for flexible joint robots - part II: Full state feedback, impedance design and experiments," in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*, vol. 3, Apr. 2004, pp. 2666–2672 Vol.3.

[30] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale, "Empirical studies in information visualization: Seven scenarios," *IEEE Trans. Vis. Comput. Graph.*, vol. 18, no. 9, pp. 1520–1536, 2012.

[31] S. Hart and L. Staveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *Advances in Psychology*, pp. 139–183, 1988.

[32] J. Brooke, "SUS: A Retrospective," *J. Usability Studies*, vol. 8, no. 2, p. 29–40, feb 2013.

[33] S. K. Tadeja, W. Rydlewicz, Y. Lu, T. Bubas, M. Rydlewicz, and P. O. Kristensson, "Measurement and Inspection of Photo-Realistic 3-D VR Models," *IEEE Computer Graphics and Applications*, vol. 41, no. 6, pp. 143–151, 2021.

[34] M. Łysakowski, K. Żywanowski, A. Banaszczyk, M. R. Nowicki, P. Skrzypczyński, and S. K. Tadeja, "Using AR and YOLOv8-Based Object Detection to Support Real-World Visual Search in Industrial Workshop: Lessons Learned from a Pilot Study," in *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 2023, pp. 154–158.

[35] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness," *Int J Aviat Psychol*, pp. 203–220, 1993.

[36] A. D. Prabaswari, C. Basumerda, and B. W. Utomo, "The mental workload analysis of staff in study program of private educational organization," *IOP Conference Series: Materials Science and Engineering*, vol. 528, no. 1, p. 012018, may 2019.

[37] K. Stanney, R. Kennedy, and J. M. Drexler, "Cybersickness is not simulator sickness," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 41, no. 2, pp. 1138–1142, 1997.

[38] S. Engeser and F. Rheinberg, "Flow, performance and moderators of challenge-skill balance," *Motivation and Emotion*, vol. 32, no. 3, pp. 158–172, Sept. 2008.

[39] M. Laakasuo, J. Palomäki, S. Abuhamdeh, O. Lappi, and B. U. Cowley, "Psychometric analysis of the flow short scale translated to Finnish," *Scientific Reports*, vol. 12, no. 1, p. 20067, Nov. 2022, number: 1 Publisher: Nature Publishing Group.

[40] F. Wilcoxon, "Individual Comparisons by Ranking Methods," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945, publisher: [International Biometric Society, Wiley].