Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 58-19 (2024) 760–765

How does the application of augmented reality affect the mental workload of human workers? A collection of preliminary results

Leonardo Maretto*, Daria Battini*, Maurizio Faccio*, Irene Granata*, Mohamad Y. Jaber**

**Department of Management and Engineering, University of Padova, Vicenza, Italy*

*** Department of Mechanical, Industrial, and Mechatronics Engineering, Toronto Metropolitan University, Toronto, Canada*

(Corresponding author e-mail: leonardo.maretto@unipd.it).

Abstract: Digitalisation and the introduction of smart manufacturing technologies are changing the shape of industrial shop floors. For manufacturing companies, it becomes imperative to assess the operational and economic advantages derived from these technologies. Simultaneously, in alignment with the Industry 5.0 paradigm, understanding the influence of digital technologies on human workers, especially those directly interacting with these technologies, is equally critical. This study focuses on augmented reality (AR) as a specific technology, conducting an experimental campaign to explore its effects on learning curves and cognitive workload. Voluntary participants engaged in learning manual tasks of varying durations and required dexterity. This study evaluates the impact of AR on the mental workload of operators, utilising two main eye-related measures: fixation duration and pupil diameter. Preliminary results indicate that augmented reality increases the cognitive workload of workers undergoing training.

Keywords: augmented reality, Industry 5.0, mental workload, digitalisation, eye tracking

1. INTRODUCTION

The novel Industry 5.0 paradigm (European Commission, 2021) emphasizes the role of emerging technologies in the manufacturing sector and their impact on workers. Augmented Reality (AR) is recognized as a promising enabling technology within the Industry 5.0 context (Maddikunta et al., 2022), particularly for its applications in learning and training. One critical aspect of AR's impact is its influence on the users' cognitive workload. Scientific literature has shown that the cognitive workload of AR use in manufacturing and manual assembly has been primarily assessed using subjective measures, with little attention given to the effects of AR during the learning process of a manual assembly task. To address this gap, this paper adopts two objective eye-related parameters, fixation duration and pupil diameter, to gauge the mental workload of individuals involved in the learning phase of a manual pick-and-place task. The rest of the paper is structured as follows: Section 2 reviews the existing literature on augmented reality and mental workload measures in manufacturing/manual assembly, Section 3 outlines the experimental setup, Section 4 discusses considerations related to measuring mental workload, Section 5 presents and discusses the experimental results and Section 6 concludes with closing remarks and future research directions.

2. LITERATURE REVIEW

This state-of-the-art section begins with general considerations on AR and mental workload, then focuses on instances of objective measures of mental (cognitive) workload in industrial manual assembly. Jeffri and Rambli (2021) conducted a comprehensive literature review on AR and mental workload across various domains. The evidence shows that, in general, augmented reality appears to have a positive impact on mental workload. Instances where AR was associated with a high mental workload were primarily linked to system usability and information presentation. Most studies utilized subjective measures, with the NASA Task Load Index (NASA-TLX) questionnaire being the most commonly employed assessment tool (Jeffri & Awang Rambli, 2021). For example, in the work of Re et al. (2016), the NASA-TLX was used to evaluate mental workload during the manual assembly of a circuit board. The questionnaire results indicated that the monitor-based AR system, which overlays instructions on a live-streamed video of the workspace, partially reduced the user's mental workload compared to traditional instruction manuals. Other studies (Marino et al., 2021) combined the NASA-TLX with another subjective measure, such as the SUS (System Usability Scale). Despite their widespread use, subjective measures may not always be well suited to assess mental workload, particularly during the learning phase of a task. Subjective measures are typically post-hoc assessments, whereas evaluating the evolution of cognitive workload during the learning process requires real-time measures. Moreover, Van Acker et al. (2020) showed that in manual assembly, subjective mental workload measures may lead to significantly different conclusions than objective measures, such as pupil size. Kim and Dey (2016) thus suggested using subjective and objective measures in assessing sensory augmentation systems. Eye-related parameters are particularly promising as objective measures because they are non-intrusive and do not significantly interfere with the subject. Marquart et al. (2015) reviewed the main eye-related measures (blinks, fixations and pupillometry) and their relationship with mental workload.

10.1016/j.ifacol.2024.09.210

However, few examples exist of using eye-related parameters as objective measures of AR-induced mental workload in manual assembly. Stork and Schubö (2010) noted that eye fixation count and duration could translate to measuring the intensity of processing information of any object in view. Wang et al. (2019) analysed the performance of Google Glass in the manual disassembly of a mobile phone, employing near-point accommodation as an eye-related measure to assess visual fatigue. Guo et al. (2022) utilised pupil diameter to measure the mental/cognitive workload of multiple subjects involved in the training phase of a telerobotic arm in space. In a similar experimental design to the one presented in this paper, Faccio et al. (2023) studied the learning curves of human workers using an AR platform, assessing mental workload with blink rate. However, this eye-related parameter was used as an overall evaluation tool without analysing its trend across multiple task repetitions. Therefore, a significant research gap exists in adopting eye-related parameters as objective measures to assess mental workload when using AR applications, particularly during the learning phase of manual tasks. Expanding on the analysis briefly introduced in Faccio et al. (2023), this study incorporates fixation duration and pupil diameter to thoroughly examine mental workload dynamics throughout training for a manual task.

3. THE EXPERIMENT

3.1 The AR platform

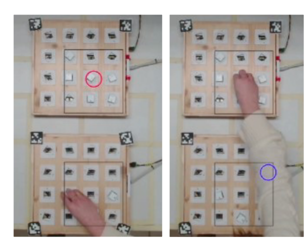
The AR technological platform was developed at the Laboratory of Industrial Plants and Logistics of the University of Padova. The system integrates a depth camera (Intel RealSense D435i) and uses the Open Pose software (Cao et al., 2021) for pose estimation. OpenPose tracks real-time movements using 25 keypoints to reproduce body pose, associating them with 2D coordinates. OpenPose generates body poses that are matched with the depth frame generated by the depth camera, utilising stereophotogrammetry for distance measurement. This ensures the generation of complete 3D coordinates. The full coordinates of the keypoints are then compared with a set of control volumes (Faccio et al., 2019, 2022), simulating manual pick-and-place or assembly tasks. The manual task was previously analysed and broken down into shorter sub-tasks, each associated with specific control volumes according to the sequence of activities. If the operator works within the correct control volume according to the sub-tasks sequence, prompts are displayed on a monitor, providing suggestions for the next step. The operator cannot proceed with the sub-task sequence until reaching the correct control volume. The system does not prompt the operator for incorrect volume positions.

3.2 Experimental design

The experiments were designed to assess the mental workload of human workers during the learning phase of a generic pick-and-place task. Two smart pallets were employed to simulate the pick-and-place activity. These pallets, equipped with sensors, are wooden boxes featuring specially designed jigs to accommodate various objects. Pick-and-place tasks involve transferring objects from one pallet to another following a predetermined sequence. The AR platform was compared with a set of traditional text-based instructions (TBI). Positions on the smart pallets were designated with numbers, with the TBI comprising a table listing pick and place positions (Fig. 1a). The table was displayed on the workstation-mounted monitor, while the object positions were reproduced on paper and affixed to the workstation. During the AR experiments, control volumes were placed around the pick and place positions. The monitor displayed a real-time stream of the smart pallets with a superimposed circle indicating the next pick or place position, which appeared only after the preceding task was completed (Fig. 1b). To test different conditions and scenarios, two types of objects were used in the experiments: spheres and cubes (Fig. 2a and 2b). Spheres required less dexterity since no orientation of the object was needed. Conversely, the cubes demanded a higher level of dexterity and overall attention: the correct orientation of the cubes was indicated by two matching dots, one on the top face of the cube and one on the jig of the smart pallet. To add complexity, certain cubes required orientation with their dot on the opposite side of the dot on the smart pallet jig, as guided by the AR system or specified in the TBI. Additionally, two task lengths were considered: 4 and 9 objects. Longer sequences with more objects required learning more combinations of pick-and-place positions. To efficiently simulate the learning process, the number of repetitions of each task was adjusted - 15 repetitions for the tasks with 4 objects and 20 repetitions for the tasks with 9. The number of repetitions struck a balance between experiment duration feasibility and generating enough data points to observe the learning curves' convergence toward a plateau (computed using the task repetition length as the main variable). Images of the experimental setup are presented in Figures 3 and 4, respectively.

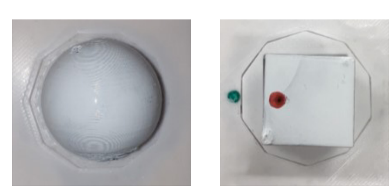
Pick	Place
3	2
4	3
2	1
1	4

a) TBI



b) AR

Figure 1. a) An example of the table shown on the monitor during the TBI test with 4 objects. b) Two screenshots of the monitor during an AR test with 9 objects. The red circle indicates the pick position and the blue circle the subsequent place position.



a) Sphere b) Cube

Figure 2. The type of objects used in laboratory experiments (the cube could have assumed either the position shown in the picture as well as a second position with the red dot on the opposite side of the green dot on the smart pallet jig)

Six male Engineering Master's students, aged 22-24, took part in this preliminary study. They formed two groups of 3. One group experimented with spheres and the other with cubes. Participants in both groups experimented with different task lengths (4 and 9 objects), using both support systems (AR and TBI). For instance, the testing plan for a participant from the sphere group included the following tests: a test with 4 spheres using TBI, a test with 4 spheres using AR, a test with 9 spheres using TBI, and a test with 9 spheres using AR. It was necessary to alter the pick-and-place sequence when switching between the TBI and AR tests to avoid learning retention.

4. MEASURING MENTAL WORKLOAD

To measure eye-related parameters and assess mental workload during the experiments, participants wore eye-tracking devices.

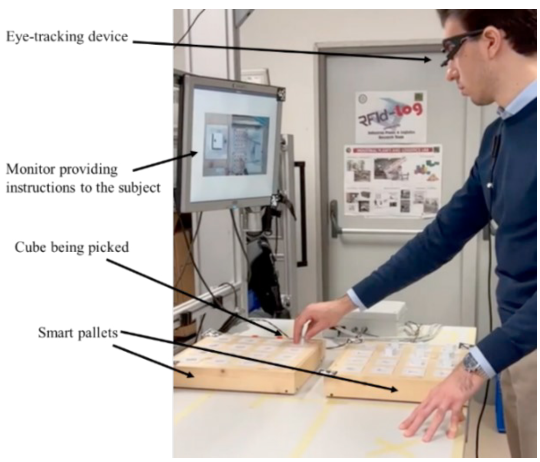


Figure 3. The experimental setup at the Laboratory for Industrial Plants and Logistics of the University of Padova

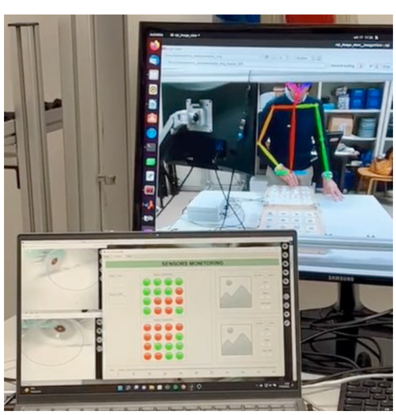


Figure 4. In the background, OpenPose generates a pose estimation of a subject. In the foreground, the output of the smart pallet sensors and the tracking of the pupil performed by the Pupil Core headset.

4.1 The eye-tracking device

The adopted eye-tracking device was Pupil Core, produced by Pupil Labs (Kassner et al., 2014). As shown in Figure 5, Pupil Core can be worn as a pair of glasses and features three cameras: a world camera capturing the participant's view and two eye cameras tracking pupil movements. The headset requires two software components: Pupil Capture for real-time data acquisition and processing, including pupil detection and calculation of eye-related measures, and Pupil Player for offline data visualisation and export.

Figure 5. The Pupil Core headset (image obtained from Pupil Labs Docs (2024)). 1) indicates the world camera, 2) the nose support, 3) the eye cameras and 4) the USB-C connector clip.

4.2 Mental workload measures

Among the eye-related parameters captured by the headset, fixation duration and pupil diameter consistently show a relationship with mental workload (Marquart et al., 2015). Fixations, defined as "pauses over informative regions of interest" (Salvucci & Goldberg, 2000), have durations typically measured in milliseconds. Fixation duration increases with increasing mental and visual effort (Marquart et al., 2015). Pupil diameter can be measured in two ways. The first (which will be referred to in this paper as 2D pupil diameter) uses each eye frame to fit a 2D ellipse (Kassner et al., 2014) and adopts pixels as the unit of measurement while the other (referred to as 3D pupil diameter) reconstructs the pupil in a 3D coordinate system using a mathematical model (Dierkes et al., 2019; Dodgson & Swirski, 2013), measuring the diameter in millimetres. Both methods offer reliable results and are supported by sound methodology, so both were employed in the experiments to assess mental workload. As with fixation duration, greater pupil diameter corresponds to increased mental workload.

4.3 Measuring mental workload across multiple task repetitions

Since the aim of this paper is to assess the mental workload of human workers during the learning phase of a manual task, the evaluation of mental workload measures across multiple repetitions of the same experiment is required. Precisely defining the start and end of each repetition is crucial. The repetition start was designated as the moment the subject completed the experiment reset, returning all objects to the initial position and preparing for a new iteration, looking at the monitor. The repetition ended with the placement of the last object in its corresponding jig. It is noteworthy that the repetition start did not coincide with the physical start of a new iteration (grasping the first object). This delay arose from necessary interactions with the experimental setup that had to be carried out by the investigator. However, during this timeframe in the "TBI scenario," instructions remained visible to the subject. This allowed for distinct cognitive learning of the sequence before physical implementation. Therefore, the repetition start was defined as the moment the subject was ready to begin, indicated by their attention to the monitor/instructions.

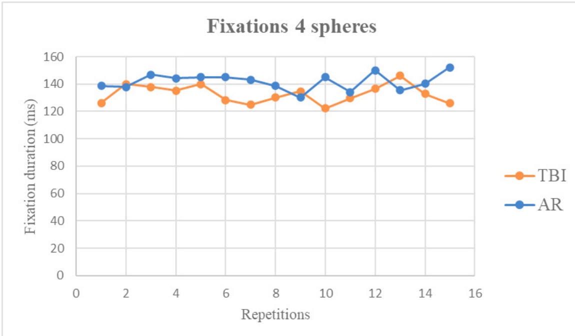


Figure 6. Fixations for the 4-sphere test

Figure 7. Fixations for the 9-sphere test

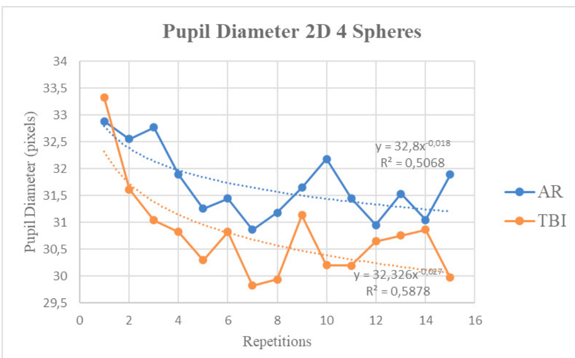


Figure 8. 2D pupil diameter for the 4-sphere test

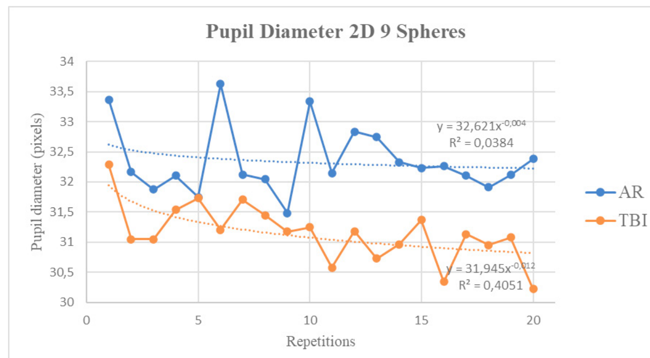


Figure 9. 2D pupil diameter for the 9-sphere test

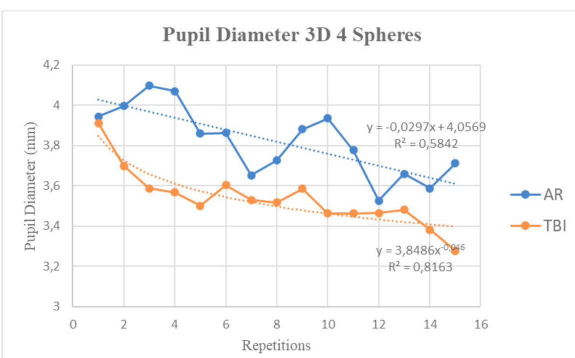


Figure 10. 3D pupil diameter for the 4-sphere test

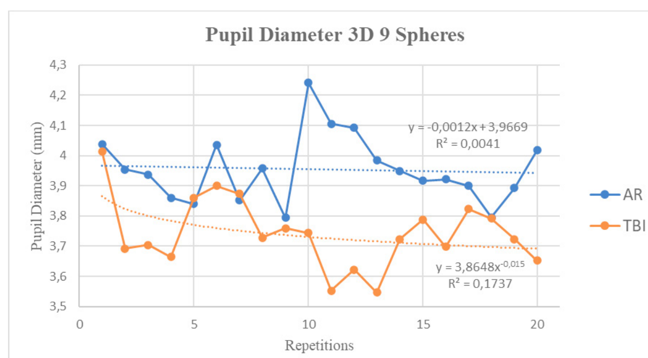


Figure 11. 3D pupil diameter for the 9-sphere test

To ensure consistency, this definition was applied even in the "AR scenario," where the operator lacked visibility of future steps (merging cognitive and physical learning aspects). Eye-related measures acquired during the experiment are stored in a comma-separated values (CSV) file generated by Pupil Player software. The high sampling rate of the headset, reaching up to 200Hz for eye cameras, yields numerous measurements for each repetition. Based on the defined start and end points for each repetition, the average value of each eye-related parameter was calculated.

5. RESULTS AND DISCUSSION

Eye-related parameters were averaged across all participants within each group. Figures 6-11 display parameter plots for the sphere tests, while figures 12-17 show those for the cube tests. Fixations exhibit no clear trend across repetition numbers, regardless of task length or required dexterity (spheres or cubes). Fixation duration is relatively stable around a constant average value. Interestingly, across all test scenarios (spheres/cubes, long/short tasks), the AR platform induces longer fixations, likely signifying increased mental workload. This finding aligns with the experiment design and tested system: AR prompts required subjects to consult the monitor following each object movement for next-step instructions. Conversely, the TBI scenario's instructions allowed participants to quickly glean the entire pick-and-place sequence from the monitor, reducing the need for frequent glances. 2D pupil diameter displays a more interesting trend. Regardless of task length, dexterity level, or support system type, 2D pupil diameter appears to decrease with increasing repetitions of the same task. Furthermore, in specific conditions (4 spheres AR, 4 spheres TBI, 4 cubes AR, 4 cubes TBI, and 9 cubes AR), the data points align well with a power curve, mirroring the earliest formulation of the learning curve in the industrial setting – Wright's learning curve (Wright, 1936).

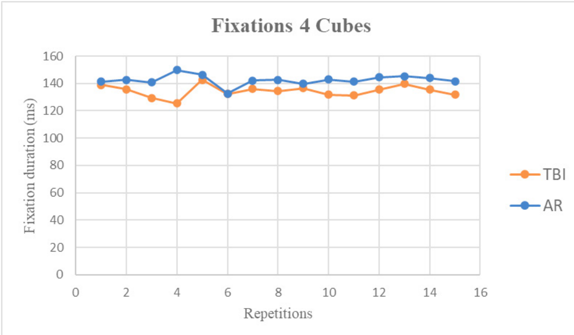


Figure 12. Fixations for the 4-cubes test

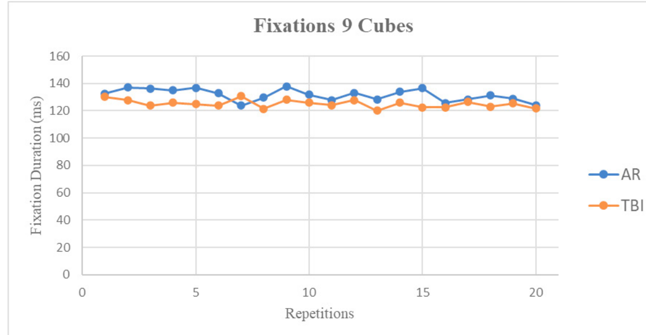


Figure 13. Fixations for the 9-cubes test

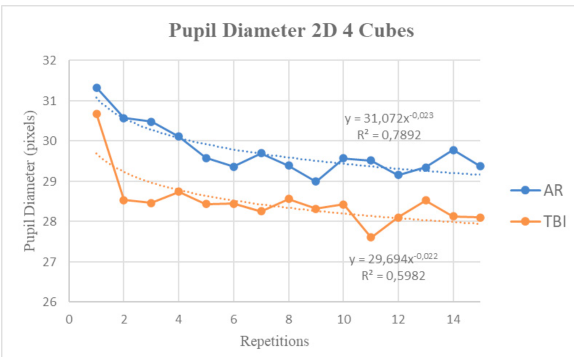


Figure 14. 2D pupil diameter for the 4-cubes test

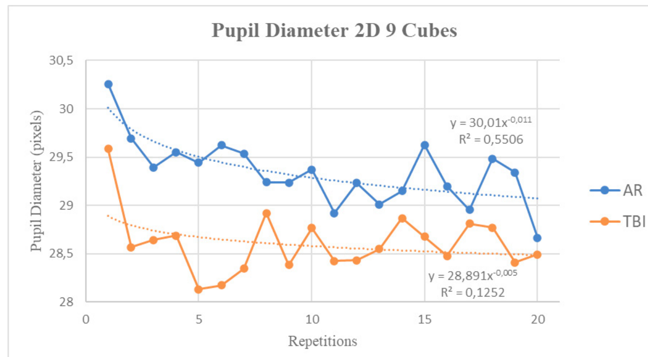


Figure 15. 2D pupil diameter for the 9-cubes test

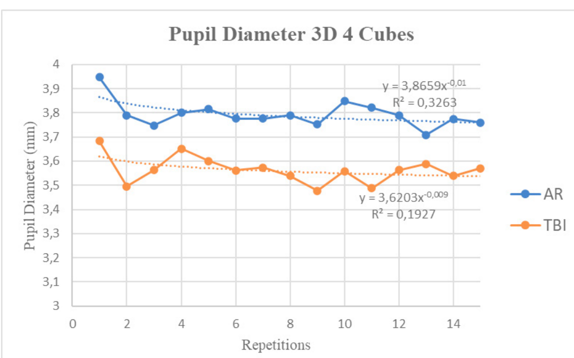


Figure 16. 3D pupil diameter for the 9-cubes test

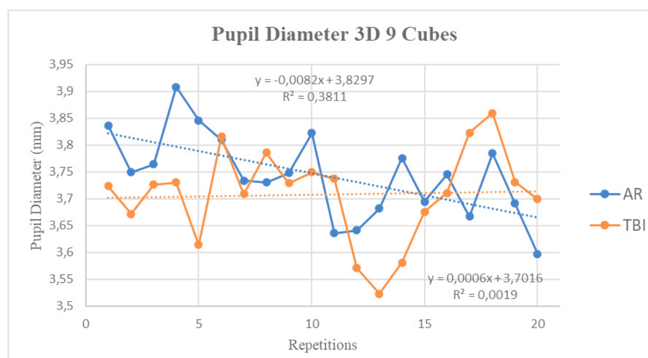


Figure 17. 3D pupil diameter for the 9-cubes test

These observations suggest that, under certain circumstances, pupil diameter, and therefore mental workload, might follow the overall manual process learning curve, as measured by repetition execution time (see Faccio et al. (2023) for a preliminary analysis of the pick-and-place sequence learning curves). Overall, the AR platform consistently induces larger pupil diameters and higher mental workload across all test scenarios. 3D pupil diameter shows a similar trend to the same 2D measure for the 4 objects test. The AR platform is responsible for larger pupil diameter but both TBI and AR follow a decreasing trend with an increase in the number of repetitions in the tests with 4 objects. On the contrary, the results are more unclear for the 9 objects tests, with the 3D pupil diameter almost constant for AR in the 9 spheres test and increasing for TBI in the 9 cubes test. This finding might indicate that, when dealing with longer tasks that require a higher dexterity level, the increase in complexity could counterbalance the learning effect on the level of the mental workload. The 3D pupil diameter findings mirror the 2D data for 4-object tests, showcasing a consistent decrease with increasing repetitions across TBI and AR platforms. This might suggest a learning effect present even with AR, despite its generally higher mental workload reflected in larger pupil diameters throughout. However, the 9-object tests present a more nuanced picture. In the 9-sphere task, the 3D pupil diameter is almost constant, while the 9-cube test shows an increase of 3D pupil diameter for TBI. This suggests that for longer tasks demanding greater dexterity, the complexity might outstrip the learning effect, potentially leading to increased mental workload despite repeated practice.

6. CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

In conclusion, AR appears to impose a higher mental workload on human workers than a traditional set of TBI during the learning phase of a manual task. In terms of eye-related parameters, TBI showed better results than AR. This is evident across most conditions, as seen in the consistently longer fixation durations and larger pupil diameters (2D and 3D) associated with the AR platform. Interestingly, both 2D and 3D pupil diameters decreased with increasing repetitions in some scenarios, mirroring the shape of a power learning curve. The main limitation of the study is the small sample size: more robust conclusions could be drawn with a larger dataset and a more rigorous statistical analysis. Additionally, investigating the influence of work environment factors like room lighting can prove valuable. Finally, future research could explore integrating a mental workload measure into a learning curve model and testing it against experimental data.

ACKNOWLEDGMENTS

The authors acknowledge the MICS (Made in Italy – Circular and Sustainable) Extended Partnership, which received funding from NextGeneration EU (Italian PNRR – M4 C2, Invest 1.3 – D.D. 1551.11-10- 2022, PE00000004). CUP MICS C93C2200528000

REFERENCES

Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., & Sheikh, Y. (2021). OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(1), 172–186.

Dierkes, K., Kassner, M., & Bulling, A. (2019). A fast approach to refraction-aware eye-model fitting and gaze prediction. *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*, 1–9.

Dodgson, N. A., & Swirski, L. (2013). *A fully-automatic, temporal approach to single camera, glint-free 3D eye model fitting*. https://www.researchgate.net/publication/264658852

European Commission, D.-G. for R. and I. B. M. , D. N. L. , P. A. (2021). *Industry 5.0 : towards a sustainable, human-centric and resilient European industry*.

Faccio, M., Ferrari, E., Galizia, F. G., Gamberi, M., & Pilati, F. (2019). Real-time assistance to manual assembly through depth camera and visual feedback. *Procedia CIRP*, *81*, 1254–1259.

Faccio, M., Granata, I., & Maretto, L. (2023). The impact of augmented reality on learning curves and mental workload: a preliminary experimental study. *2023 9th International Conference on E-Society, e-Learning and e-Technologies*, 35–39.

Faccio, M., Guidolin, M., Maretto, L., & Reggiani, M. (2022). A Framework for a Closed Loop Control System of a Human Operator in a Manual Workstation. *Proceedings of the Summer School Francesco Turco*.

Guo, Y., Freer, D., Deligianni, F., & Yang, G.-Z. (2022). Eye-Tracking for Performance Evaluation and Workload Estimation in Space Telerobotic Training. *IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS*, *52*(1).

Jeffri, N. F. S., & Awang Rambli, D. R. (2021). A review of augmented reality systems and their effects on mental workload and task performance. In *Heliyon* (Vol. 7, Issue 3). Elsevier Ltd.

Kassner, M., Patera, W., & Bulling, A. (2014). Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, 1151–1160.

Kim, S. J., & Dey, A. K. (2016). Augmenting human senses to improve the user experience in cars: applying augmented reality and haptics approaches to reduce cognitive distances. *Multimedia Tools and Applications*, *75*(16), 9587–9607.

Maddikunta, P. K. R., Pham, Q. V., B, P., Deepa, N., Dev, K., Gadekallu, T. R., Ruby, R., & Liyanage, M. (2022). Industry 5.0: A survey on enabling technologies and potential applications. In *Journal of Industrial Information Integration* (Vol. 26). Elsevier B.V.

Marino, E., Barbieri, L., Colacino, B., Fleri, A. K., & Bruno, F. (2021). An Augmented Reality inspection tool to support workers in Industry 4.0 environments. *Computers in Industry*, *127*.

Marquart, G., Cabrall, C., & de Winter, J. (2015). Review of Eye-related Measures of Drivers' Mental Workload. *Procedia Manufacturing*, *3*, 2854–2861.

Pupil Labs Docs. (2024, February).

Re, G. M., Oliver, J., & Bordegoni, M. (2016). Impact of monitor-based augmented reality for on-site industrial manual operations. *Cognition, Technology and Work*, *18*(2), 379–392.

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the Symposium on Eye Tracking Research & Applications - ETRA '00*, 71–78.

Stork, S., & Schubö, A. (2010). Human cognition in manual assembly: Theories and applications. *Advanced Engineering Informatics*, *24*(3), 320–328.

Van Acker, B. B., Bombeke, K., Durnez, W., Parmentier, D. D., Mateus, J. C., Biondi, A., Saldien, J., & Vlerick, P. (2020). Mobile pupillometry in manual assembly: A pilot study exploring the wearability and external validity of a renowned mental workload lab measure. *International Journal of Industrial Ergonomics*, *75*.

Wang, C. H., Tsai, N. H., Lu, J. M., & Wang, M. J. J. (2019). Usability evaluation of an instructional application based on Google Glass for mobile phone disassembly tasks. *Applied Ergonomics*, *77*, 58–69.

Wright, T. P. (1936). Factors Affecting the Cost of Airplanes. *Journal of the Aeronautical Sciences*, *3*(4), 122–128.