Visual Assembly Guidance under Cognitive Load: Insights from an Eye-Tracking Study

Stefan LECHNER[1], David KOSTOLANI[1], Jan MIKUNI[2], Sebastian SCHLUND[1]

[1] Department of Human-Machine Interaction, TU Wien, Austria, Theresianumgasse 27, 1040 Vienna
[2] Cognitive Sciences Hub, University of Vienna, Austria, Kolingasse 14-16, 1090 Vienna

Abstract: Spatial Augmented Reality (SAR) in assembly has been shown to improve productivity by augmenting work instructions. Due to this reason, SAR is frequently used to support workers with disabilities in performing new assembly tasks. Although this form of assembly guidance promises to improve learning, little is known about the optimal design of instructions when working under cognitive load. In this paper, we investigate differences in the visual perception of assembly instructions projected by a SAR system in the presence of cognitive load. Through an exploratory study with five participants, we induced cognitive load through a concurrent counting task while assembling large-sized toy building blocks. Results show that under high cognitive load, individuals followed visual cues less frequently, with increased eye saccade velocity.

Keywords: Spatial Augmented Reality, Eye-Tracking, Cognitive Load, Visual Assembly Guidance, Assembly Instructions, Saccades

1. Introduction

In the pursuit of enhancing accessibility for disabled individuals, Spatial Augmented Reality (SAR) is frequently used to augment work instructions directly in their field of view. This form of visual support promises to enhance work productivity by decreasing error rates and cycle times (Rupprecht et al. 2022), while also improving overall cognitive workload. As SAR can be used to break down complex tasks into more manageable work steps, many claim that visual assembly guidance can also improve learning new tasks, thus aiding the integration of persons with special needs into complex assembly (Heinz-Jakobs et al. 2022; Baechler et al. 2016).

While performing an unknown task or learning, one is exposed to and processes new information, resulting in cognitive load. This, in turn, can impede other concurrent visual or motoric movements, hindering the worker from their task. This study explores the interplay between cognitive load and moving projections, specifically within the context of an assembly task. In our experiment, a worker builds a human-sized tower with the help of projections. The projections are generated through a ceiling-mounted projector and an electrical pivotable mirror from Dynamic Projection Institute (2024). By examining eye movements during this task, we aim to uncover how cognitive load influences information comprehension during dynamic projections.

To understand where the participants look in a task or how the eyes behave in certain conditions, a device that captures those movements is needed. Eye-tracking was achieved through the Pupil Core (Kassner et al. 2014) of the company Pupil Labs, a head-mounted device that features two infrared (IR) cameras with IR illumination for the eyes and a world camera that captures the field of view. The pupil movement is translated with the intrinsic and extrinsic parameters of the cameras to an XY-position of the world camera image stream. With this information, further conclusions can be drawn, namely fixations, gaze, and saccades. Fixations represent moments of focused attention, gaze indicates the ongoing direction of sight, and saccades are rapid eye movements.

2. Background

Integration of people with disabilities has served as one of the main objectives for more inclusion in the industry. Historically, assembly has emerged as one of the main areas with high potential for the integration of workers with disabilities. Within the assembly of complex manufacturing parts, workers usually follow paper or computer-based instructions (Funk et al. 2015). However, this way of displaying instructions can result in a split attention effect, which in turn negatively influences learning and results in a high cognitive load (Vanneste et al. 2020).

The use of SAR for the provision of work instructions has been shown to mitigate the split attention effect by integrating visual instructions into the work environment. Compared to other forms of augmented reality, SAR does not require any device that needs to be carried by the worker. Instead, SAR relies on a projector to display visuals in the environment. This way, the field of view corresponds to the natural view of the user, and eye accommodation and the use of peripheral vision are easier to achieve (Zhou et al. 2011). Therefore, SAR has been shown to result in higher productivity than head-mounted displays (Büttner et al. 2016).

Due to these positive effects, prior works have investigated SAR for the provision of step-by-step instruction for workers with disabilities. One of the driving research factors is that SAR results in lower mental load and higher usability, which allows people with disabilities to get integrated into more complex assembly processes (Aksu et al. 2019). Further studies have shown that SAR increases productivity and decreases error rates compared to displaying instructions via PC screens (Funk et al. 2015). Similarly, Mark et al. (2021) found that people with disabilities prefer SAR, and it allows them to work with a higher degree of accuracy. However, while these studies have shown that SAR can improve productivity indicators for workers with disabilities, little is known about how to best design instructions for situations with increased cognitive load, such as when learning a new task or in the case of complex assembly. To our knowledge, no prior works have deployed eye-tracking for a better understanding of visual attention in projection-aided assembly.

3. Study

In our observation study, we engaged five participants of working age who have a university background. Prior to the work task, we calibrated the eye-tracker for each participant and established a baseline for eye-tracking while they sat for approximately two minutes. Subsequently, participants were directed to the task, during which controlled projections were manipulated by a study supervisor. Thus, the participants could focus solely on the task. The projections indicated which building bricks were to be picked up and highlighted their designated assembly positions. Together, they constituted a 34-step process. Following the completion of the assembly, we introduced cognitive load through a counting task and then instructed the participants to reverse-build the tower. The counting task began with participants randomly selecting a high number and then counting backwards at a self-selected interval. This ensured they chose an appropriate level of cognitive distraction such that they could still perform the assembly. This sequence was repeated for all participants, allowing us to investigate the impact of cognitive load on projection tracking. The videos from the eye-tracker were post-processed and intersubjectively analysed with Pupil Player (Kassner et al. 2014) based on the moving projections. Furthermore, the resulting annotated data were analysed for the saccade movement.

4. Results

The resulting videos showed that while the participants were not under cognitive load, they exhibited a higher degree of engagement in tracking the moving projection with their eyes compared to when they were under cognitive load. Two examples are shown in Figure 1, each illustrating the eye movement pattern. Despite this observable trend, the extracted features failed to provide a robust distinction, rendering it insufficient to draw conclusive insights based solely on these. A driving constraint was the small field of view of the Pupil Core camera, which led to data loss, especially because both fixations and gaze focused more on the local hand area apart from deviations when the projection moved. Consequently, accurate assignment of the projection's start was sometimes hindered, especially in instances where participants directed their gaze too much downward.

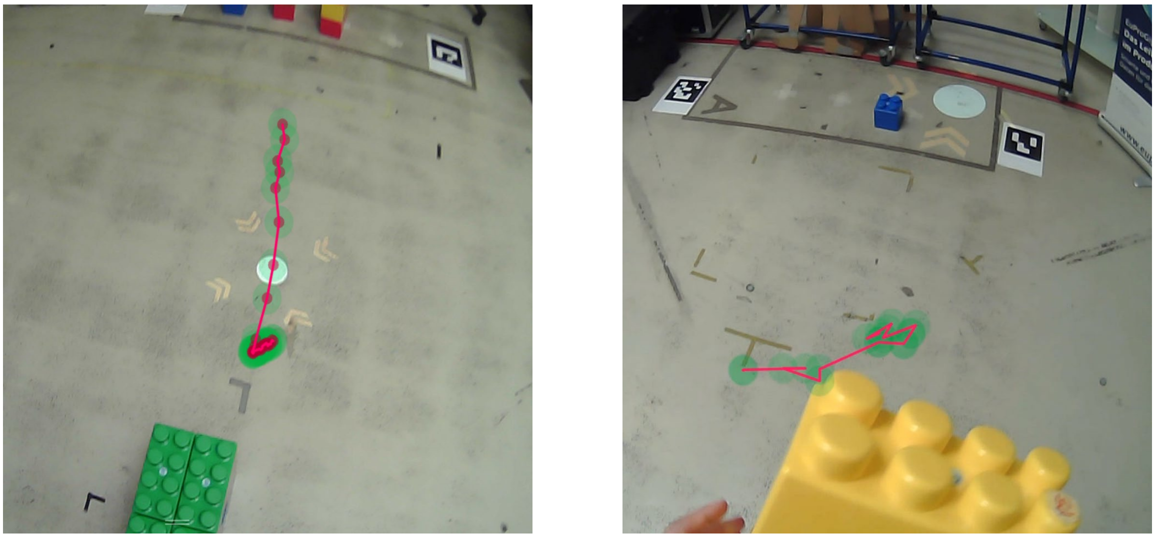

***Figure 1**: World camera view of eye-tracking device and visualisation of gaze locations per video frame, highlighted in red. The projected moving white circle denotes both the direction and position. The left image represents without, and the right under cognitive load situation.*

The change of the saccades (saccadic velocity) in the x and y directions shows a much clearer distinction. Figure 2. shows the extracted data of the x direction. The left side shows the movement pattern while working without cognitive load and the right side with cognitive load. Notably, a discernible contrast is observed, revealing a denser movement sequence with higher average amplitude during cognitive load. The rise in saccadic velocity correlates with heightened cognitive workload in all participants. This phenomenon can be attributed to the fact that, during a viewing task, individuals endeavour to gather more information within the same timeframe by executing faster saccadic movements (Bodala et al. 2014).

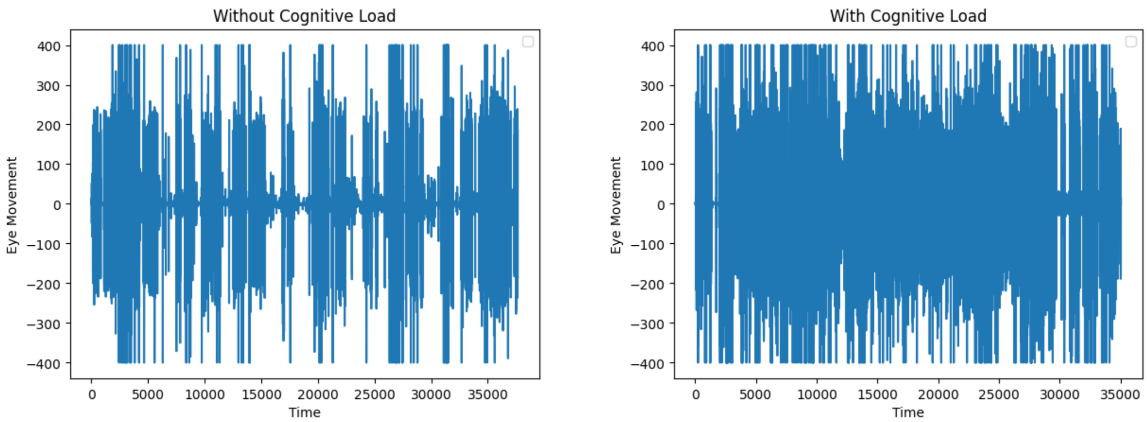

***Figure 2**: Saccadic velocity in the x-direction, with the left side representing instances of no cognitive load and the right side indicating situations involving cognitive load.*

5. Discussion

In summary, our findings highlight a correlation between cognitive load and participants' engagement in tracking a moving projection, indicative of concurrent multitasking challenges. Liu (1996) reveals that the impact of visual scanning on task performance and subjective workload is pronounced in dual-task conditions. Concluding from the threaded cognition theory by Salvucci & Taatgen (2008), the interplay of the moving projection and continuous cognitive load introduces a bottleneck, providing insight into the worsening tracking performance.

Furthermore, the observed increase in saccade velocity under cognitive load suggests challenges in focusing on dynamic projected information, potentially compromising the user's ability to process and comprehend displayed content. Saccades are needed to understand one's surrounding environment. With an increase in saccades, ambient perception can be influenced. To underscore this point, we draw a parallel between tracking a moving projection and navigating an ever-changing environment. The escalated saccade velocity arises from participants' efforts to comprehend the environment swiftly, but the constant changes place additional strain on the understanding of the environment. Being guided through a large workspace ultimately results in some orientation loss, emphasising the intricate relationship between cognitive load, saccades, and the effective processing of dynamic visual information e.g. tracking a moving projection.

Based on our study, we found that it would be best to mitigate the information density in a moving projection. It still can be rich when the projection stops but should be reasonably simple while moving. It seems that the size of the projection also plays an important role. We propose to use large and slow projections compared to small and fast-moving ones. A moving projection that traces a trajectory onto the workspace is an intriguing projection type to consider. This approach can help users keep track of information and prevent them from missing important details. As a result, managing threaded cognition should become more feasible.

Noteworthy limitations are the sample size of the study and the eye-tracking device that had a rather narrow field of view. Although every participant showed the same cues, a larger study with a more compatible device should be performed to further study the reported effects. Another limitation was the lack of randomisation in our exploratory study. A within-subject design with a randomly assigned cognitive load for the same task activity could be performed. Consequently, this should result in a stronger and more reliable relationship between cognitive load and eye-tracking data. In the future, more research is needed to understand how cognitive load and multitasking influence a person who uses projection-based guiding systems in large-scale environments.

6. Conclusion

In this study, we observed an influence of cognitive load on both visual tracking and an increased saccadic velocity during projection-assisted assembly. The conclusion highlights key areas for future research, exploring the impact of complex information in moving projections, verifying behaviour across individuals with different cognitive abilities, assessing spatial awareness risks, and investigating various projection types. Acknowledging the exploratory nature of our study, we invite further inquiry into these areas, emphasising the need to refine visual guiding systems and explore challenges faced by cognitively impaired individuals in the development of assistive SAR systems. Overall, our findings contribute to understanding the implications of cognitive load on SAR systems and their effectiveness, guiding future research in this evolving field.

7. Literature

Aksu V, Jenderny S, Kroll B & Röcker C (2019). A digital assistance system providing step-by-step support for people with disabilities in production tasks. In Advances in Usability, User Experience and Assistive Technology: Proceedings of the AHFE 2018 International Conferences on Usability & User Experience and Human Factors and Assistive Technology, Held on July 21–25, 2018, in Loews Sapphire Falls Resort at Universal Studios, Orlando, Florida, USA 9 (pp. 775-785). Springer International Publishing

Baechler L, Baechler A., Funk M, Autenrieth S, Kruell G, Hoerz T & Heidenreich T (2016). The use and impact of an assistance system for supporting participation in employment for individuals with cognitive disabilities. In Computers Helping People with Special Needs: 15th International Conference, ICCHP 2016, Linz, Austria, July 13-15, 2016, Proceedings, Part II 15 (pp. 329-332). Springer International Publishing.

Bodala IP, Ke Y, Mir H, Thakor NV & Al-Nashash H (2014, August). Cognitive workload estimation due to vague visual stimuli using saccadic eye movements. In 2014 36th annual international conference of the IEEE engineering in medicine and biology society (pp. 2993-2996). IEEE.

Büttner S, Funk M, Sand O & Röcker C (2016). Using head-mounted displays and in-situ projection for assistive systems: A comparison. In Proceedings of the 9th ACM international conference on pervasive technologies related to assistive environments (pp. 1-8).

Dynamic Projection Institute GmbH (2024). https://www.dynamicprojection.com, Accessed January 2, 2024

Funk M, Mayer S & Schmidt A (2015, October). Using in-situ projection to support cognitively impaired workers at the workplace. In Proceedings of the 17th international ACM SIGACCESS conference on Computers & accessibility (pp. 185-192).

Heinz-Jakobs M, Große-Coosmann A & Röcker C (2022, September). Promoting Inclusive Work with Digital Assistance Systems: Experiences of Cognitively Disabled Workers with In-Situ Assembly Support. In 2022 IEEE Global Humanitarian Technology Conference (GHTC) (pp. 377-384). IEEE.

Kassner M, Patera W & Bulling A (2014). Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication (pp. 1151-1160).

Liu Y (1996). Quantitative assessment of effects of visual scanning on concurrent task performance. Ergonomics, 39(3), 382-399.

Mark BG, Rauch E & Matt DT (2021). The application of digital worker assistance systems to support workers with disabilities in assembly processes. Procedia CIRP, 103, 243-249.

Rupprecht P, Kueffner-McCauley H, Trimmel M, Hornacek M & Schlund S (2022). Advanced Adaptive Spatial Augmented Reality utilizing Dynamic in-situ Projection in Industrial Site Assembly. Procedia CIRP, 107, 937-942.

Salvucci DD & Taatgen NA (2008). Threaded cognition: an integrated theory of concurrent multitasking. Psychological review, 115(1), 101.

Vanneste P, Huang Y, Park JY, Cornillie F, Decloedt B & Van den Noortgate W (2020). Cognitive support for assembly operations by means of augmented reality: an exploratory study. International Journal of Human-Computer Studies, 143, 102480.

Zhou J, Lee I, Thomas B, Menassa R, Farrant A & Sansome A (2011, December). Applying spatial augmented reality to facilitate in-situ support for automotive spot welding inspection. In Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry (pp. 195-200).

Acknowledgement: This work was supported by Vienna Chamber of Labour as a part of the A2I project (Nr. 6-483, Digitalisierungsfonds Arbeit 4.0).

Gesellschaft für Arbeitswissenschaft e.V.

Arbeitswissenschaft in-the-loop

Mensch-Technologie-Integration und ihre Auswirkung auf Mensch, Arbeit und Arbeitsgestaltung

70. Kongress der
Gesellschaft für Arbeitswissenschaft e.V.

Institut für Arbeitswissenschaft und
Technologiemanagement IAT
Universität Stuttgart

In Zusammenarbeit mit dem Fraunhofer-Institut für
Arbeitswirtschaft und Organisation IAO

06. – 08. März 2024

GfA-Press

Bericht zum 70. Arbeitswissenschaftlichen Kongress vom 06. – 08. März 2024

Institut für Arbeitswissenschaft und Technologiemanagement (IAT), Universität Stuttgart
In Zusammenarbeit mit: Fraunhofer-Institut für Arbeitswirtschaft und Organisation (IAO), Stuttgart

Herausgegeben von der Gesellschaft für Arbeitswissenschaft e.V.
Sankt Augustin: GfA-Press, 2024
ISBN 978-3-936804-34-8

NE: Gesellschaft für Arbeitswissenschaft: Jahresdokumentation

Als Manuskript zusammengestellt. Diese Jahresdokumentation ist nur in der Geschäftsstelle (s. u.) erhältlich.

Geschäftsstelle der GfA
Simone John, Tel.: +49 (0)30 1300-13003, Alte Heerstraße 111, D-53757 Sankt Augustin
info@gesellschaft-fuer-arbeitswissenschaft.de · www.gesellschaft-fuer-arbeitswissenschaft.de

Screen design und Umsetzung
© 2024 fröse multimedia, Frank Fröse,
office@internetkundenservice.de, www.internetkundenservice.de