p-ISSN: 2184-044X e-ISSN: 2184-1489 ISBN: 978-989-35728-0-1 © 2024
https://doi.org/10.36315/2024v2end083

AUGMENTED REALITY GLASSES IN CLASS – BLESSING OR CURSE

Verena Schmuck, Teresa Spiess, & Christian Ploder
Management Center Innsbruck (Austria)

Abstract

This paper focuses on the current state of augmented reality (AR) usage in university teaching and the impact of AR on student workload. Universities need to engage with new technologies and developments at an early stage. An empirical experiment was conducted with two independent groups to investigate the effects of Microsoft HoloLens 2 on students' subjective workload when solving a Sudoku puzzle. The experimental group used the Microsoft HoloLens 2, and the control group used a tablet. The NASA Task Load Index measured workload and its dimensions. The experiment was based on previous research on AR in higher education and Cognitive Load Theory. HoloLens users' workload was significantly higher (average 70.356/100) than tablet users (average 55.896/100). HoloLens users rated the physical demands higher in the individual components of the workload, and their subjective performance was significantly worse than that of the control group. Furthermore, the workload was mainly related to the HoloLens 2 experience, but students' affinity for technology did not influence these scores. No differences in gender-based workload could be found when using HoloLens 2.

Keywords: *Augmented reality, NASA TLX, teaching, practical training.*

1. Introduction

Colleges and universities are recognized as catalysts for innovation and are essential in driving advancements in various fields. The influence of digitalization extends to education and research activities at universities. Hence, it is crucial for these institutions to proactively engage with emerging technologies at an early stage, develop concepts, and formulate potential use cases, as emphasized by Lilligreen and Wiebel (2019). Augmented reality (AR) is one such technology already applied in research, despite its conceptual existence since 1970, showcasing constant evolution (Knoll & Stieglitz, 2022). Various applications and scenarios can be developed and implemented using AR (Tolba, Elarif & Taha, 2022, p. 53). While AR has been extensively explored in the literature, a research gap still needs to be explored, particularly concerning AR glasses and their correlation with student workload. The workload of students significantly influences their well-being and, consequently, their academic performance (Salmela-Aro & Read, 2017). The escalating relevance of digitalization and its associated impact on student workload is gaining attention in the research community (Van De Velde et al., 2021, p. 114). Therefore, investigating this aspect concerning the Microsoft HoloLens 2 is deemed crucial.

The paper begins by defining the research objectives and questions in section 2. They were followed by the theoretical background, which describes the concept of AR and workload. Then, the research method is explained, split up into hypotheses, a sample of the experiment, data collection and empirical setting, questionnaires, and the concepts of affinity of technology and the NASA Task Load Index (NASA TLX). Then, the result will be discussed and wrapped up by a research conclusion.

2. Objectives

On the one hand, the theoretical part of this paper aims to explain the concept of AR and workload to the reader. Furthermore, an empirical experiment is set up based on the literature. In the empirical part, the following research question is examined by an empirical experiment: “What are the effects of using the Microsoft HoloLens 2 on students' subjective workload when solving a Sudoku puzzle?” Further questions that can be derived from the research question and are investigated in this context: (1) How does the workload differ between the experimental group (using the HoloLens) and the control group (using a tablet)? (2) Is there a connection between students' affinity for technology and their subjective workload when using the HoloLens, and (3) Is there a correlation between objective performance (score achieved) and the subjectively assessed performance of students when using the HoloLens?

3. Theoretical background

Azuma (1997) defines AR as a technology that enhances or overlays the real world with virtual objects. The foundation of AR is the physical reality, which is then enriched with additional elements and AR components (Saxena & Verma, 2022). In addition to supplementing reality, AR can remove virtual objects or conceal portions of the actual environment. These virtual objects contain information that is invisible in reality and help users accomplish tasks in the real world. Users experience a sense of coexistence between the real world and augmented objects, intensifying their perception of reality and fostering increased interaction with it (Azuma, 1997). Azuma's (1997) definition does not confine itself to a specific technology. Still, it outlines certain properties that characterize AR systems: The application combines the real world with the virtual world, functions interactively in real-time, and registration takes place in 3D. Hart (2006) defines workload as an individual's effort to complete a task. The effort has implications for an individual's physical and mental well-being, with increased effort potentially leading to symptoms of exhaustion, heightened emotional stress, diminished performance, or increased cognitive strain (Hart & Wickens, 1990). However, workload is influenced by the individual's physical, mental, and emotional state, necessitating consideration of factors such as fatigue, stress, or illness in this context (Hart, 2006). Measuring workload involves various methods, often combining subjective and objective data to assess its value (Said et al., 2020, p. 9). One commonly used survey instrument for this purpose is the multi-level National Aeronautics and Space Administration-Task Load Index (NASA-TLX) developed by Hart and Staveland (1988).

4. Methodology and measurements

This section will explain the methodology used for the data collection and give a clear overview of the experiment. The used Sudoku is shown in Figure 1, and the measurement metrics are provided.

A literature review was conducted at the beginning of the research. The keywords "Augmented Reality", "Augmented Reality in Education", Augmented Reality AND Cognitive Load", Workload AND Augmented Reality", "Workload in Education" were used, and the term "Augmented Reality" was also replaced by "HoloLens" and "Smart Glasses". EBESCO, PubMed BibSearch, Springer Link, and Google Scholar were used. In addition, an experiment in combination with several questionnaires was carried out for the empirical part of the work.

An empirical experiment was conducted to answer the research question. A two-group design (experimental group vs. control group) was developed to measure the workload using the technologies (Döring & Bortz, 2016, p. 194). The present experiment can be categorized as a laboratory study in a controlled environment. However, this also means that the results could change in a typical environment and everyday situations, examinations, or lecture situations, which harms external validity (Döring & Bortz, 2016, p. 206). Only students who were studying for a Bachelor's or Master's degree were allowed to take part in this study. A total of 31 students, encompassing both Bachelor's and Master's students, took part. The participant breakdown included 17 females and 14 males. Half of the sample specialized in IT studies or pursued IT/computer science degrees, while the remaining half pursued various fields of study. Some participants had prior experience with augmented reality. The experiments were conducted from May 5 to 7, 2023, and took place in the library and study area of the Management Center Innsbruck.

First, the participants answered the first questionnaire for each group. The first questionnaire, part 1 of both groups, contains nine items on affinity for technology at the beginning. The HoloLens group was then asked questions about their experiences with AR and the HoloLens, while the control group was asked the same questions about their experiences with the tablet. In addition, previous experiences with Sudoku were surveyed, with the respective answer options for the different experiences matching between the groups. Then, the students were given the task description and an introduction to the fundamental rules of Sudoku. Following an overview of the technique, the participants had approximately five minutes to familiarize themselves with the respective devices, practice writing, and address any unanswered questions. Subsequently, the actual task was presented. Each participant was challenged to solve a Sudoku puzzle classified as "easy" level. Group 1 utilized the HoloLens 2, while Group 2 completed the task using a tablet.

The objective for the test subjects was to achieve the same Sudoku solution within a 15-minute timeframe or to solve the Sudoku as fast as possible. Participants earned one point for each correctly filled box, with a maximum possible score of 45 points for a fully completed puzzle. To incentivize successful Sudoku completion, the participant with the highest or all correct boxes, irrespective of group, received a cash voucher. After finishing the Sudoku task, all study participants received the standardized NASA Task Load Index questionnaire, designed to assess subjective workload.

p-ISSN: 2184-044X e-ISSN: 2184-1489 ISBN: 978-989-35728-0-1 © 2024
https://doi.org/10.36315/2024v2end083

Subsequently, the participants answered part 2 of the questionnaire specific to their respective groups. This included personal attitudes toward Augmented Reality (AR) and HoloLens 2 for the HoloLens group and attitudes toward tablets for the tablet group. Additionally, participants from both groups indicated the last time they had played Sudoku. Finally, demographic questions were posed to gather information about the test subjects and draw conclusions about the sample. These questions covered age, gender, academic degree (Bachelor's or Master's), field of study, IT specialization, and whether participants wore glasses. Post-experiment, the NASA TLX was utilized again to determine subjective workload. In the second part of the questionnaire for both groups, assessments were collected regarding the experiences and feelings while using the respective technology in the task, aiming to conclude potential influencing factors. All questions were asked in German, and the experimenter was present during the experiment to resolve any content ambiguities immediately and answer experiment participants' questions. Afterward, results were counted, and points were calculated, counting all correctly filled recognizable numbers.

The participants were tasked with solving a Sudoku, a logic puzzle that requires filling in missing fields. A Sudoku comprises 81 squares, with some fields pre-filled, and the remaining empty fields must be completed to solve the puzzle. The puzzle's difficulty is determined by the number of initially provided numbers, with more challenging puzzles having fewer given numbers. Each Sudoku is individually generated (Delahaye, 2006).

The Sudoku presented in Figure 1 is the puzzle participants had to solve in the experiment using either the HoloLens 2 or the Samsung Galaxy Tab S6. The difficulty level of this Sudoku was classified as "easy," and it contained 49 correct fields that could be filled in.

Figure 1. Sudoku.

	9	1	5	7	3	8	2	
	5	2	8		6	9	4	7
	7	6		4				
	8			3				
2					5	7		1
5	4					6	8	
			3	2		4		8
	2		1		8			
7							1	

Other independent variables about the Sudoku task could also be gathered. Firstly, the number of points achieved could be derived from the correctly completed boxes. Secondly, time was recorded, or a time limit of 15 minutes was imposed.

Affinity for technology interaction (ATI) characterizes an individual’s tendency to engage with technical systems actively or to avoid them. The score reflects a person's technology utilization and proficiency in using it. It is assumed that the higher this indicator is for an individual, the more efficiently and productively that person utilizes technology. The term "technological systems" in this concept encompasses applications such as apps, software, and all digital devices. Digital devices include cell phones, computers, televisions, and car navigation systems. ATI (Franke, Attig & Wessel, 2019). The ATI score is measured by a standardized questionnaire with nine items on a Likert scale.

The National Aeronautics and Space Administration - Task Load Index (NASA-TLX) is a multidimensional rating scale designed to measure the subjective workload experienced by individuals during or after a task. The questionnaire comprises two parts. In the first part, the dimensions of mental, physical, and temporal demands, as well as performance, effort, and frustration, are examined based on the definition provided by Hart and Staveland (1988).

5. Results and discussion

The evaluations revealed that the subjective workload of HoloLens students was higher than the control group, who were given the same task and used a tablet to solve it. In particular, significantly higher values were recorded in the performance and physical effort categories. The tablet group assessed their performance considerably higher than the HoloLens group, while the HoloLens group perceived the physical effort as significantly more significant than the control group. Furthermore, this study did not find a correlation between NASA TLX and ATI when using the HoloLens. However, the findings from a survey conducted by Attig et al. (2018) suggest that students' ATI scores can still play a crucial role in avoiding the high workload associated with the HoloLens. By adjusting the training intensity based on the ATI score, providing more or less intensive training with this technology becomes possible, thus improving usability. The study implies that, depending on the ATI score, customized training and support are essential to prevent users from feeling overwhelmed, ultimately leading to improved usability.

This observation is consistent with the finding that students with more excellent experience using these technologies exerted less effort. This could be attributed to technologically adept students approaching the task more quickly and confidently. The connection between high self-confidence and lower workload, as outlined by Said et al. (2020, p. 8), further supports the notion that experience with technology influences user workload. Consequently, the higher workload experienced with the HoloLens may be attributed to the group's lower familiarity with the device than the group using the tablet. A correlation was observed between the attained score and the self-assessed performance in the HoloLens group. The higher the score, the better the students perceived their performance, aligning with findings from previous studies in medical applications. This relationship is also evident in medical studies, such as the one conducted by Said et al. (2020), which employed NASA-TLX to gauge the workload in patient monitoring. Notably, a correlation was identified between test performance and the total and partial values of NASA-TLX. Instances of good performance were associated with lower NASA-TLX scores, indicating a lower perceived workload. Participants tended to rate their performance higher when test results were better, suggesting that increased training and experience with a task can reduce subjective workload and enhance performance. As emphasized earlier, experience and training influence results (Said et al., 2020, p. 8).

This study found a correlation between HoloLens experience and subjective perceived workload. Contrary to expectations, the study revealed that more excellent experience with the HoloLens was associated with a higher perceived workload. While this may seem counterintuitive, a similar finding was reported by Hu, Lu, Tan, and Lomanto (2016, p. 1742), who discovered that simulation training improved estimated performance and individual workload ratings. This suggests that training with the HoloLens may enhance the experience, subsequently impacting subjective workload perceptions. The study also explored whether gender has an impact on workload. No discernible differences were identified in the perceived workload when using the HoloLens based on gender. This outcome could suggest that both male and female students had similar prior experiences with the technology. No significant differences in mental workload were observed between HoloLens and tablet use in this study. However, existing research in education and AR has demonstrated that technology can assist learners in engaging with the material through interactive elements and information on position and shapes (Rios, Hincapié, Caponio, Mercado & González Mendívil, 2011). Studies have indicated that high interactivity in AR applications can decrease cognitive load compared to traditional methods, such as paper-based learning (Klinker, Tabarani, Wiesche & Krcmar, 2020, p. 388). For instance, medical students using mobile AR for anatomy learning achieved better performance and experienced lower cognitive load than their colleagues (Küçük, Kapakin & Göktaş, 2016, p. 419). AR learning materials, with their ability to present complex information compactly and symbolically, focus attention on specific points, thereby reducing mental strain and lowering the load on working memory (İbili, 2019, p. 52). While the study found no significant differences in mental and cognitive demands between the two groups engaged in solving Sudoku puzzles, it is noteworthy that the technology's impact on student workload is influenced by various factors, including the design of AR applications and the interactivity of elements (Lilligreen & Wiebel, 2019).

In summary, the study indicates that AR is suitable for learning, as it enhances motivation and when implemented effectively in the classroom, facilitates learning by providing more interactive opportunities for students, thereby influencing retention (İbili, 2019). While the potential for AR in education is evident, challenges remain, including the need to adapt educational programs and increased application options for students and teachers, and the technology's current limitations make permanent integration into teaching impractical. Nonetheless, as AR devices become more affordable and accessible, there is potential for broader adoption in educational settings (Lilligreen & Wiebel, 2019).

6. Conclusion and future research

To sum up, the use of HoloLens 2 increases the subjective workload in comparison to the use of a tablet. However, workload dimensions should always be examined individually, and the applications' previous experience and user-friendliness should be considered. For further experiments, it would be necessary to enlarge the number of participants and look for a higher diversity.

p-ISSN: 2184-044X e-ISSN: 2184-1489 ISBN: 978-989-35728-0-1 © 2024
https://doi.org/10.36315/2024v2end083

References

Attig, C., Mach, S., Wessel, D., Franke, T., Schmalfuß, F., & Krems, J. F. (2018). Technikaffinität als Ressource für die Arbeit in Industrie 4.0. *Aw&I Conference, 3*. https://doi.org/10.14464/awic.v3i0.251

Azuma, R. T. (1997). A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, *6*(4), 355-385. https://doi.org/10.1162/pres.1997.6.4.355

Delahaye, J.-P. (2006). The science behind Sudoku. *Scientific American, 294*(6), 80-87. https://doi.org/10.1038/scientificamerican0606-80

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5th ed.). Berlin, Heidelberg: Springer Berlin Heidelberg.

Franke, T., Attig, C. & Wessel, D. (2019). A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale. *International Journal of Human-Computer Interaction, 35*(6), 456-467. https://doi.org/10.1080/10447318.2018.1456150

Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50*(9), 904-908. https://doi.org/10.1177/154193120605000909

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Empirical and Theoretical Research Results. *Advances in Psychology, 52*, 139-183. https://doi.org/10.1016/S0166-4115(08)62386-9

Hart, S. G., & Wickens, C. D. (1990). Workload Assessment and Prediction. In H. R. Booher (Ed.), *Manprint* (pp. 257–296). Dordrecht: Springer Netherlands.

Hu, J. S. L., Lu, J., Tan, W. B., & Lomanto, D. (2016). Training improves laparoscopic task performance and decreases operator workload. *Surgical Endoscopy, 30*(5), 1742-1746. https://doi.org/10.1007/s00464-015-4410-8

İbili, E. (2019). Effect of augmented reality environments on cognitive load: pedagogical effect, instructional design, motivation, and interaction interfaces. *International Journal of Progressive Education, 15*(5), 42-57. https://doi.org/10.29329/ijpe.2019.212.4

Klinker, K., Tabarani, G., Wiesche, M., & Krcmar, H. (2021). Augmented Reality für Aufgaben mit hoher kognitiver Belastung. In M. Wiesche, I. M. Welpe, H. Remmers, & H. Krcmar (Eds.), *Informationsmanagement und digitale Transformation. Systematische Entwicklung von Dienstleistungsinnovationen* (pp. 371–392). Wiesbaden: Springer Fachmedien Wiesbaden.

Knoll, M., & Stieglitz, S. (2022). Augmented Reality and Virtual Reality – Einsatz im Kontext von Arbeit, Forschung und Lehre. *HMD Praxis der Wirtschaftsinformatik, 59*(1), 6-22. https://doi.org/10.1365/s40702-022-00840-5

Küçük, S., Kapakin, S., & Göktaş, Y. (2016). Learning anatomy via mobile augmented reality: Effects on achievement and cognitive load. *Anatomical Sciences Education, 9*(5), 411-421.

Lilligreen, G., & Wiebel, A. (2019). Augmented Reality in Vorlesung und Übung: Lehre und Interaktion neu gedacht. In T. Barton, C. Müller, & C. Seel (Eds.), Angewandte Wirtschaftsinformatik. *Hochschulen in Zeiten der Digitalisierung* (pp. 221–238). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-26618-9_14

Rios, H., Hincapié, M., Caponio, A., Mercado, E., & González Mendívil, E. (2011). Augmented Reality: An Advantageous Option for Complex Training and Maintenance Operations in Aeronautic Related Processes. In R. Shumaker (Ed.), *Lecture Notes in Computer Science. Virtual and Mixed Reality - New Trends* (Vol. 6773, pp. 87-96). Berlin, Heidelberg: Springer Berlin Heidelberg.

Salmela-Aro, K. & Read, S. (2017). Study engagement and burnout profiles among Finnish higher education students. *Burnout Research, 7*, 21-28. https://doi.org/10.1016/j.burn.2017.11.001

Saxena, D., & Verma, J. K. (2022). Recreating Reality: Classification of Computer-Assisted Environments. In J. K. Verma & S. Paul (Eds.), *Studies in Computational Intelligence. Advances in Augmented Reality and Virtual Reality* (Vol. 998, pp. 3-9). Singapore: Springer Singapore.

Tolba, R., Elarif, T., & Taha, Z. (2022). Augmented reality in technology-enhanced learning: Systematic review 2011-2021. *International Journal of Intelligent Computing and Information Sciences, 22*(1), 1-16. https://doi.org/10.21608/ijicis.2022.97513.1121

Van De Velde, S., Buffel, V., Bracke, P., Van Hal, G., Somogyi, N. M., Willems, B., &Wouters. (2021). The COVID-19 International Student Well-being Study. *Scandinavian Journal of Public Health, 49*(1), 114–122. https://doi.org/10.1177/1403494820981186

Said, S., Gozdzik, M., Roche, T. R., Braun, J., Rössler, J., Kaserer, A., Spahn, D. R., Nothiger, C. B., & Tscholl, D. W. (2020). Validation of the Raw National Aeronautics and Space Administration Task Load Index (NASA-TLX) Questionnaire to assess perceived workload in patient monitoring tasks: Pooled analysis study using mixed models. *Journal of Medical Internet Research, 22*(9), e19472.