

Video Analysis of Behavioral Patterns During Prolonged Work in VR

Verena Biener^{1*} Forouzan Farzinnejad¹ Rinaldo Schuster¹ Seyedmasih Tabaei¹ Leon Lindlein¹
 Jinghui Hu² Negar Nouri¹ John J. Dudley² Per Ola Kristensson² Jörg Müller³
 Jens Grubert^{1†}

¹Coburg University of Applied Sciences and Arts, Germany

²University of Cambridge, United Kingdom

³University of Bayreuth, Germany



a) Eating or Drinking



b) Rubbing Eyes or Face



c) Physical World Activities

Figure 1: Actions that were observed during the video annotation process. (a) Eating or drinking; (b) Rubbing eyes or face; (c) Interacting with the physical world objects, like phones or papers. These actions were observed less frequently in VR.

ABSTRACT

VR has recently been promoted as a tool for knowledge workers and studies have shown that it has the potential to improve knowledge work. However, studies on its prolonged use have been scarce. A prior study compared working in VR for one week to working in a physical environment, focusing on performance measures and subjective feedback. However, a nuanced understanding and comparison of participants' behavior in VR and the physical environment is still missing. To this end, we analyzed video material made available from this previously conducted experiment, carried out over a working week, and present our findings on comparing the behavior of participants while working in VR and in a physical environment.

Keywords: virtual reality, video-analysis, productivity work, prolonged use, office work, future of work

1 INTRODUCTION

The opportunities for Virtual Reality (VR) as a medium for knowledge work, as opposed to entertainment, are beginning to be explored [1, 3, 7]. VR has the potential to enhance the working experience through multiple mechanisms, such as exploiting large virtual displays [3, 6], personalizing the work-environment [7], or delivering enhanced interactivity [1]. However, for anyone who has worn a current-generation VR headset for an extended period of time, the notion of completing even a single work day wearing a head-mounted display (HMD) could be a daunting prospect.

The demands of knowledge work, in contrast to VR in gaming or leisure, suggest that workers may either willingly or unwillingly spend extended periods wearing an HMD without the welcome distraction of entertainment. Recognition of this fact motivates research seeking to understand how knowledge workers respond to extended use of VR [5, 8]. So far the longest study was conducted by Biener et al. [2] in which participants were completing a full work week in

VR while doing their normal work tasks. This was then compared to a week in which participants worked in a comparable physical setup. They found that the VR condition delivered significantly worse ratings across measures of task load, frustration, negative affect, anxiety, eye strain, system usability, flow, productivity, wellbeing and simulator sickness. Nevertheless, some of the reported measures slightly improved over the five days. While conducting this study, Biener et al. [2] amassed a rich dataset of over 1,400 hours of video, capturing participant behavior in both conditions throughout both weeks, but this was not analysed in the original paper. We used this dataset and analyzed videos of one day per condition for all 16 participants to report on their behavior as they respond to the experience of working in VR as compared to working in the physical environment. To our knowledge, this is the first study of its kind and delivers substantial insight into the, as yet, undocumented behavior of users in a novel VR work setting.

2 METHODOLOGY

The goal of this work was to closely analyze the behavior of 16 participants (6 female, 10 male, mean age = 29.31 years, $sd = 5.52$) during an experiment involving prolonged use of VR in a work setting. The participants were observed wearing HMDs during their work for five continuous days (VR condition) and then for another five days without an HMD (PHYSICAL condition), as described in Biener et al. [2]. Each day, they worked for eight hours with a 45 minute lunch break after four hours. The videos we use here were recorded during that study [2] using a webcam. We analyzed the participants' behavior, by involving six people in watching and annotating interesting behavior in the videos (Annotator A, B, C, D, E, and F). Through an iterative process, annotator A and B created a codebook for labelling behavior in VR. The codebook for the PHYSICAL videos was derived as closely as possible from the VR codebook, allowing a comparison of the behavior from VR with the standard behavior in PHYSICAL. The other annotators were then trained by annotator A and B and samples of annotations were compared between annotators. To obtain more consistency, all videos of one participant were only worked on by one annotator.

Due to the high demand on time for labeling one video (about one hour for processing one hour of video material), we report only on the comparison of the fifth day in VR and the first day of PHYSICAL.

*e-mail: verena.biener@hs-coburg.de

†e-mail: jens.grubert@hs-coburg.de

This was chosen, because participants would be most familiar with the HMD on their last day and because participants were already familiar with working with a standard desktop setup, so we do not expect a change of behavior over time for the PHYSICAL videos.

3 RESULTS

We analyzed the occurrence of different categories of events by comparing DAY 5 of VR to DAY 1 of PHYSICAL using a repeated measures ANOVA with the independent variable INTERFACE (VR, PHYSICAL) and TIME (MORNING, AFTERNOON), similar as in prior work [2]. To get the average per hour, we divided the total number and duration of events during each time-period (MORNING, AFTERNOON) by the duration of this time-period, which was usually around four hours and 22 minutes (four hours of work plus half of the 45 minute break). Whenever the sphericity assumption was violated, we used Greenhouse-Geisser correction to ensure the robustness of the ANOVA [4]. We applied Bonferroni-correction to all post-hoc tests involving multiple comparisons.

Screen Time: We calculated screen time in VR by considering all times in which participants were wearing the HMD, not including times in which they were taking the HMD halfway off, using the controller, having problems with the keyboard tracking or were reading, writing or using their phone. In PHYSICAL we labelled all times in which the participants were facing the screen, also not including times in which they were reading, writing or using the phone. This approach could not be reliably used for VR, because it is not clear from the videos if participants were looking at the virtual monitor. Comparing VR to PHYSICAL showed no main effect of INTERFACE, suggesting that the time participants spent working was not significantly different between conditions.

Standing up: We labelled all parts where participants were standing up or sitting down. Comparing VR to PHYSICAL showed no significant difference in the number of standing events, however, we found a significant influence of TIME ($F(1, 15) = 66.97, p < 0.001, \eta_p^2 = 0.82$) on the duration of standing per hour and no interaction effect between INTERFACE and TIME. Post-hoc tests indicated that both in VR and PHYSICAL participants stood significantly less ($p < 0.001$) during the AFTERNOON (VR: $m = 734.08 \text{ sec}, sd = 438.32$; PHYSICAL: $m = 750.38 \text{ sec}, sd = 594.27$) than during the MORNING (VR: $m = 3605.99 \text{ sec}, sd = 1459.87$; PHYSICAL: $m = 3244.79 \text{ sec}, sd = 1561.97$).

Eating or Drinking: This category describes all parts of the video in which participants put drinks or food into their mouth, such as displayed in Fig. 1 (a), not including longer periods of chewing. Comparing PHYSICAL to VR, we found that the INTERFACE had a significant influence on the number of such events per hour ($F(1, 15) = 7.97, p = 0.013, \eta_p^2 = 0.45$), such that in PHYSICAL ($m = 3.19, sd = 2.67$) participants were drinking or eating more often than in VR ($m = 1.79, sd = 1.31$). However, we could not find a significant influence on the time spent eating or drinking per hour. This suggests that participants took longer to eat or drink in VR which could be caused by them being more cautious.

Rubbing Eyes or Face: This category was used whenever participants were rubbing their eyes or face, as illustrated in Fig. 1 (b). When comparing VR to PHYSICAL, we found a significant main effect of INTERFACE ($F(1, 15) = 5.6, p = 0.032, \eta_p^2 = 0.27$) on the number of such actions, indicating that participants rubbed their eyes and faces more in PHYSICAL ($m = 2.87, sd = 2.25$) than VR ($m = 1.68, sd = 1.44$). We also found significant main effects of TIME ($F(1, 15) = 7.64, p = 0.014, \eta_p^2 = 0.34$) indicating more such actions in the MORNING ($m = 2.75, sd = 2.36$) compared to the AFTERNOON ($m = 1.8, sd = 1.36$). However, post-hoc tests did not confirm this for VR and PHYSICAL individually. We also did not find a significant effect of INTERFACE on the total time spent on such actions.

Physical World Activities: We were also interested in how often and for how long participants concerned themselves with things outside of the virtual world, as depicted in Fig. 1 (c). Therefore, this category includes events where the annotators believe the participants were reading or writing something outside of VR, using a smartphone, or otherwise peeking under the HMD. Comparing VR to PHYSICAL showed a significant main effect of INTERFACE ($F(1, 15) = 7.669, p = 0.014, \eta_p^2 = 0.34$) such that there are significantly more such actions in PHYSICAL ($m = 6.88, sd = 8.40$) than in VR ($m = 3.28, sd = 4.97$), but no significant differences have been found regarding the total time spent on such actions.

4 DISCUSSION AND CONCLUSION

Analyzing the videos showed that some actions were more common in PHYSICAL as compared to VR such as consuming food, rubbing faces and eyes or interacting with the physical world. However, no significant difference could be found for the total duration of such events per hour. An explanation for this could be that it takes more effort to perform such actions in VR and therefore participants do them less often, but for prolonged times, or more cautiously. Also in PHYSICAL some of these actions could happen involuntarily. For other actions (screen time, standing up), no significant difference could be found between PHYSICAL and VR, on the contrary, similar behavior was detected such as less time spent standing in the afternoon.

These results already provide interesting insights into the behavior of users while working in VR for a much longer timescale than the duration of many VR user studies so far. However, the videos recorded in the prior study [2] can provide much more insights into user behavior than presented in the scope of this paper, such as the changes in behavior over the period of a full work week and detailed observations that are only accessible through a video analysis. Therefore, we are planning to analyze and report on such more in-depth findings in future work.

REFERENCES

- [1] V. Biener, T. Gesslein, D. Schneider, F. Kawala, A. Otte, P. O. Kristensson, M. Pahud, E. Ofek, C. Campos, M. Kljun, et al. Povrpoint: Authoring presentations in mobile virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 28(5):2069–2079, 2022.
- [2] V. Biener, S. Kalamkar, N. Nouri, E. Ofek, M. Pahud, J. J. Dudley, J. Hu, P. O. Kristensson, M. Weerasinghe, K. Č. Pucihar, et al. Quantifying the effects of working in vr for one week. *IEEE Transactions on Visualization and Computer Graphics*, 28(11):3810–3820, 2022.
- [3] V. Biener, D. Schneider, T. Gesslein, A. Otte, B. Kuth, P. O. Kristensson, E. Ofek, M. Pahud, and J. Grubert. Breaking the screen: Interaction across touchscreen boundaries in virtual reality for mobile knowledge workers. *arXiv preprint arXiv:2008.04559*, 2020.
- [4] M. J. Blanca, J. Arnau, F. J. García-Castro, R. Alarcón, and R. Bono. Non-normal data in repeated measures anova: Impact on type i error and power. *Psicothema*, 35(1):21–29, 2023.
- [5] J. Guo, D. Weng, H. Fang, Z. Zhang, J. Ping, Y. Liu, and Y. Wang. Exploring the differences of visual discomfort caused by long-term immersion between virtual environments and physical environments. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 443–452. IEEE, 2020.
- [6] M. McGill, A. Kehoe, E. Freeman, and S. Brewster. Expanding the bounds of seated virtual workspaces. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 27(3):1–40, 2020.
- [7] A. Ruvimova, J. Kim, T. Fritz, M. Hancock, and D. C. Shepherd. "transport me away": Fostering flow in open offices through virtual reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2020.
- [8] R. Shen, D. Weng, S. Chen, J. Guo, and H. Fang. Mental fatigue of long-term office tasks in virtual environment. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 124–127. IEEE, 2019.