Experiences with User Studies in Augmented Reality

Marc Satkowski
Interactive Media Lab Dresden
Technische Universität Dresden
Dresden, Germany
msatkowski@acm.org

Wolfgang Büschel
Interactive Media Lab Dresden
Technische Universität Dresden
Dresden, Germany
bueschel@acm.org

Raimund Dachselt[*][†]
Interactive Media Lab Dresden
Technische Universität Dresden
Dresden, Germany
dachselt@acm.org

CCS CONCEPTS

• **Human-centered computing** → **User studies**; *Mixed / Augmented Reality*.

1 INTRODUCTION

The research field of augmented reality (AR) is of increasing popularity, as seen, among others, in several recently published surveys [5, 6, 8–10, 12]. To produce further advancements in AR, it is not only necessary to create new systems or applications, but also to evaluate them. One important aspect in regards to the evaluation is the general understanding of how users experience a given AR application, which can also be seen by the increased number of papers focusing on this topic [12] that were published in the last years. With the steadily growing understanding and development of AR in general, it is only a matter of time until AR devices make the leap into the consumer market where such an in-depth user understanding is even more essential. Thus, a better understanding of factors that could influence the design and results of user experience studies can help us to make them more robust and dependable in the future.

In this position paper, we describe three challenges which researchers face while designing and conducting AR users studies. We encountered these challenges in our past and current research (see Fig. 1), including papers that focus on perceptual studies of visualizations [4, 14], interaction studies [3], and studies exploring the use of AR applications [1, 2] and their design spaces [11].

2 CHALLENGES IN AR USER STUDIES

User studies are a fundamental tool in HCI research. However, the design process can be rather difficult and challenging, especially for AR systems. Among those challenges, we want to name the following: **(1)** more studies should be conducted in in-the-wild scenarios [6, 9, 12, 14], while a long-term user experience should be of concern [7], **(2)** generating insights can be difficult [12], even with data collected under the best possible conditions [7], and **(3)** therefore a need for "new evaluation methods that could capture more accurately the user experience in AR" [10]. In general, we will address two main questions that are connected to those challenges but also derived by our own experience:

- Q1 How can we ensure that the participants of a user study can easily give answers to various questions?
- Q2 How can experimenters make sure that the participants solve the given tasks at hand in the correct manner and that the study prototype works as intended?

Based on those two questions, we will discuss the challenges of the input capabilities of head-mounted display (HMD) AR devices, the communication between the participants and the experimenter, and the needed user study system setup in regards to the two previous challenges. Here we want to note, that we will only focus on optical see-through HMDs like the HoloLens v1 and the HoloLens 2, which we used in our own studies. Therefore, some of the mentioned problems have to be assessed differently for other device types, like video-see through devices.

2.1 Input Capabilities of AR

Some parts of the user experience rely on the feeling how responsive and easy to use a system is. However, the input capabilities of AR devices are rather challenging, both in regards to the general interaction, but also for completing specific study tasks and answering questionnaires. The default way to fill out questionnaires, which we also often utilize in our AR studies, is to use pen and paper or a desktop computer. However, this procedure forces the participants to leave the designated study space and sometimes to remove the HMD as well. To continue the study, it can be necessary to re-calibrate or restart the AR application which can introduce small alterations to the setup. To minimize this problem, it would be possible to use the AR specific input modalities (in regards to the Microsoft HoloLens), like free-hand gestures, to fill out such questionnaires. Yet, this possibility has different problems associated with it:

- P1 Most of the invited participants in our user studies are inexperienced with AR and its input capabilities. Even an extended training session for free-hand gestures can only give them a shallow understanding, while increasing the overall study duration.
- P2 The general recognition of free-hand gestures is error prone, due to possible tracking errors. This can be amplified by certain subtleties of a gesture (e.g., hand has to be in the field of view of the HMD camera) that inexperienced participants are not aware of. This can lead to not or wrongly recognized gestures.
- P3 Greater fluctuations between answer quality and task completion time can appear for different tasks and participants while they use free-hand gestures as an input modality. Again, this can be linked to inexperience or possible recognition errors.

[*] Also with, Centre for Tactile Internet with Human-in-the-Loop (CeTI), Technische Universität Dresden.
[†] Also with, Cluster of Excellence Physics of Life, Technische Universität Dresden.

Figure 1: Different images from a few user study setups. (A) [14] A participants sits in front of a real-life algae reactor which is part of an industrial production plant. He has to analyze the shown visualization, while interacting via a Microsoft Clicker in his right hand. Tasks were answered orally. (B) [3] In this study, we explored how mobile devices like a smartphone can be used to pan and zoom in 3D data spaces in AR. Here, the user holds the smartphone in her right hand and can use the spatial movement of the device to alter the view on the point cloud. Additionally, to unify their coordinate systems, the Microsoft HoloLens as well as the smartphone are equipped with tracking markers for a motion tracking system. (C) [11] Two persons, each wearing a HoloLens, analyze different data visualizations with a combination of four tablets. In this work, we presented a prototype that combines several mobile devices, HoloLenses, a motion tracking system, as well as an additional server application on a desktop PC.

- P4 Not all tasks and input types are suited for free-hand gestures, like text input.
- P5 An extended use of free-hand gestures can also lead to greater exhaustion, due to additional body, arm, or hand movement.

In general, the usage of free-hand gestures allows the input needed for the answering of questionnaires, but seems rather slow, error prone, or too complicated for the participants. This is why different approaches can be used and explored to minimize the use of free-hand gestures for filling out questionnaires or the general interaction with an AR application:

- A1 A reduction of the introduced interaction set can help to minimize the training needs and possible recognition errors. This is not always possible, especially for studies that explore complete AR systems and applications. However, for studies that focus on perception or design choices, it can be a suitable solution.
- A2 The usage of other input devices can leverage on other already known input modalities, like the Microsoft Clicker as a single mouse button (see Fig. 1a). Again, this requires a quite simple interaction set, and in regards to the Clicker, can only be based on pointing. We used this, for example, in [14].
- A3 Mobile devices, like smartphones or tablets, are another device type that can be used for interaction in combination with AR (see Fig. 1b). In general, we are already quite familiar how to write text or fill out forms with such devices. Yet, the introduction of such devices brings other problems with it, which will be further described in subsection 2.3.
- A4 Participants could give feedback orally, like in an interview. We used this approach in [14]. This can be easily realized, but makes it harder for the participants to think about the question, to understand the question and its answer due to high noise levels in the environment, and to select an answer from a predefined set of answer (e.g., rating scales).

2.2 Participant-Experimenter Communication

Another essential part of user experience investigations is the ability to talk with participants about the presented systems or designs, to verify and log what the participants do, and to communicate with the participants about certain parts of a system. However, an AR HMD allows only one person to observe the virtual content, which generally prevents the experimenter from observing the study participants. To address this, it is possible to extend the study application by different approaches:

- A1 The AR application could be accompanied by an additional experimenter client on, e.g., a desktop computer, which logs specific activities and events made by the participant or the system. We used this, for example, in [14]. This allows the investigator to react to given predefined events but only gives a coarse understanding of the current user behaviour in the AR application and provides no visual feedback.
- A2 The experimenter could use the streaming capabilities of the AR HMD to see exactly what the participants currently do and see. We used this in our study [11]. However, this approach has several drawbacks, like a possible latency due to network issues, decreased performance on the AR device due to increased computational requirement (e.g., the HoloLens 2 will be limited to 30 fps with no antialiasing), and possible offsets in the position of virtual objects.
- A3 The experimenter could also wear an AR HMD to see exactly what the participants see. We used this, for example, in [1]. However, this approach requires at least two HMDs and could introduce other problems that we further describe in subsection 2.3.

In general, those additions allow for a better communication between the experimenter and the participant, but also have relative high costs in regards of development expenses and performance. Lastly, with this increased need of preparation and implementation it takes far longer to create a study application.

2.3 Complexity of AR Study Setups and Multi Device Applications

AR applications are inherently more complex in comparison to desktop or smartphone applications, which makes it in general harder to analyze and study the overall user experience. Such systems do not exist in isolation and are therefore connected with several other devices and objects all around them in the near environment. Thus, it is only natural, that AR HMDs have to coexist with several other systems (see Fig. 1c), like **(1)** motion tracking systems for information about the position and rotation of objects, **(2)** other mobile devices like smartphones and tablets with potential web applications, **(3)** stationary devices like desktops or larger vertical displays, **(4)** other AR devices that should be used in the same virtual space, or **(5)** other intelligent (e.g., washing machine, machines in industry 4.0) and normal everyday objects. However, this results in additional problems:

- P1 With the increase of possible devices and objects that should be connected to an AR application, the costs for the development of such a system increases as well. This is mainly caused by the need for additional testing, implementation of fallback mechanisms, synchronizing the local coordinate systems, and state synchronicity between each object.
- P2 Since AR applications aim to be adopted in real-life scenarios, such applications should also be able to be used and studied in those. However, the availability of the needed infrastructure and interference by other present devices and networks could pose additional problems in such environments.

One approach to reduce the complexity of such systems is to provide and use different frameworks that encapsulate specific functionalities. Those can contain system behaviours (e.g., MRTK), visualizations (e.g., u2Vis [13]), server communication, or interaction management.

3 CONCLUSION

In this position paper we, presented three different challenges for AR user studies, specific problems linked to those, and possible approaches to counter them. However, many of the mentioned problems are not easy to solve and the proposed solutions could introduce even more problems. To be able to study how users perceive and experience such AR systems, it is important to be mindful about those challenges and to address them accordingly. Such a robust system is especially important for user experience studies, since not only individual interactions, design concepts, or visualizations are investigated, but a whole application will be experienced by the users. This is why possible side effects, like a decreased frame rate or faster exhaustion, can be quite critical and detrimental to the overall user experience. Since we only described a small subset of challenges connected to head-mounted AR devices, we want to note that there exist many more possible factors that can influence the users and their performance. We hope that this position paper serves as a starting point for further discussions of the challenges in evaluating user experiences in Mixed Reality.

ACKNOWLEDGMENTS

This work was funded in part by "Deutsche Forschungsgemeinschaft" (DFG, German Research Foundation) under grant number 319919706/RTG2323 "Conducive Design of Cyber-Physical Production Systems", under project number 389792660 as part of TRR 248 (see https://perspicuous-computing.science), under Germany's Excellence Strategy - EXC-2050/1 - Project ID 390696704 - Cluster of Excellence "Centre for Tactile Internet with Human-in-the-Loop" (CeTI) of Technische Universität Dresden, and under Germany's Excellence Strategy - EXC-2068 - 390729961 - Cluster of Excellence "Physics of Life" of Technische Universität Dresden.

REFERENCES

[1] Wolfgang Büschel, Anke Lehmann, and Raimund Dachselt. 2021. MIRIA: A Mixed Reality Toolkit for the In-Situ Visualization and Analysis of Spatio-Temporal Interaction Data. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, New York, NY, USA. https://doi.org/10.1145/3411764.3445651

[2] Wolfgang Büschel, Annett Mitschick, and Raimund Dachselt. 2018. Here and Now. In *Proceedings of the 2018 Conference on Human Information Interaction&Retrieval - CHIIR '18*. ACM Press, New York, New York, USA, 171–180. https://doi.org/10.1145/3176349.3176384

[3] Wolfgang Büschel, Annett Mitschick, Thomas Meyer, and Raimund Dachselt. 2019. Investigating smartphone-based pan and zoom in 3D data spaces in augmented reality. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2019*. Association for Computing Machinery, Inc. https://doi.org/10.1145/3338286.3340113

[4] Wolfgang Buschel, Stefan Vogt, and Raimund Dachselt. 2019. Augmented Reality Graph Visualizations. *IEEE Computer Graphics and Applications* 39, 3 (2019), 29–40. https://doi.org/10.1109/MCG.2019.2897927

[5] Luís Fernando de Souza Cardoso, Flávia Cristina Martins Queiroz Mariano, and Ezequiel Roberto Zorzal. 2020. A survey of industrial augmented reality. *Computers & Industrial Engineering* 139 (2020), 106159. https://doi.org/10.1016/j.cie.2019.106159

[6] Arindam Dey, Mark Billinghurst, Robert W. Lindeman, and J. Edward Swan. 2018. A systematic review of 10 Years of Augmented Reality usability studies: 2005 to 2014. *Frontiers Robotics AI* 5, APR (2018). https://doi.org/10.3389/frobt.2018.00037

[7] Barrett Ens, Benjamin Bach, Maxime Cordeil, Ulrich Engelke, Marcos Serrano, Wesley Willett, Arnaud Prouzeau, Christoph Anthes, Wolfgang Büschel, Cody Dunne, Tim Dwyer, Jens Grubert, Jason H Haga, Nurit Kirshenbaum, Dylan Kobayashi, Tica Lin, Monsurat Olaosebikan, Fabian Pointecker, David Saffo, Nazmus Saquib, Dieter Schmalstieg, Danielle Albers Szafir, Matthew Whitlock, and Yalong Yang. 2021. Grand Challenges in Immersive Analytics. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, New York, NY, USA. https://doi.org/10.1145/3411764.3446866

[8] Austin Erickson, Kangsoo Kim, Gerd Bruder, and Gregory Welch. 2020. A Review of Visual Perception Research in Optical See-Through Augmented Reality. *ICAT-EGVE* (2020), 1–9. https://doi.org/10.2312/egve.20201256

[9] Adrien Fonnet and Yannick Prie. 2019. Survey of Immersive Analytics. *IEEE Transactions on Visualization and Computer Graphics* (2019), 1–22. https://doi.org/10.1109/tvcg.2019.2929033

[10] Kangsoo Kim, Mark Billinghurst, Gerd Bruder, Henry Been Lirn Duh, and Gregory F. Welch. 2018. Revisiting trends in augmented reality research: A review of the 2nd Decade of ISMAR (2008-2017). *IEEE Transactions on Visualization and Computer Graphics* 24, 11 (2018), 2947–2962. https://doi.org/10.1109/TVCG.2018.2868591

[11] Ricardo Langner, Marc Satkowski, Wolfgang Büschel, and Raimund Dachselt. 2021. MARVIS: Combining Mobile Devices and Augmented Reality for Visual Data Analysis. In *Proceedings of the 2021 ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA. https://doi.org/10.1145/3411764.3445593

[12] Leonel Merino, Magdalena Schwarzl, Matthias Kraus, Michael Sedlmair, Dieter Schmalstieg, and Daniel Weiskopf. 2020. Evaluating Mixed and Augmented Reality: A Systematic Literature Review (2009-2019). In *Proceedings - 2020 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2020*. Institute of Electrical and Electronics Engineers Inc., 438–451. https://doi.org/10.1109/ISMAR50242.2020.00069 arXiv:2010.05988

[13] Patrick Reipschlager, Tamara Flemisch, and Raimund Dachselt. 2021. Personal Augmented Reality for Information Visualization on Large Interactive Displays. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 1182–1192. https://doi.org/10.1109/TVCG.2020.3030460 arXiv:2009.03237 [cs.HC]

[14] Marc Satkowski and Raimund Dachselt. 2021. Investigating the Impact of Real-World Environments on the Perception of 2D Visualizations in Augmented Reality. In *Proceedings of the 2021 ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA. https://doi.org/10.1145/3411764.3445330