Affordance Guided User Elicitation of Interaction Concepts for Unimodal Gaze Control of Potential Holographic 3D UIs in Automotive Applications

Maryia Kazhura*
Technische Universität Berlin
Volkswagen AG

Bea Vorhof*
Universität zu Lübeck

André Calero Valdez*
Universität zu Lübeck

Figure 1: HoloLens 2 implementation of affordance guided gaze interaction concepts for eight exemplary interaction tasks, obtained from the user elicitation survey. The tasks were designed for use with a 240 × 135 mm large holographic 3D display, emulated with a HoloLens 2.

ABSTRACT

Identifying novel but still usable and intuitive interaction methods for specific technologies and Cross Reality (CR) applications remains a challenge for developers and user interface designers. The process can become even more complex when the target hardware is still in prototype stages of development and does not support usability tests in early design iterations. Using an affordance guided user centered elicitation survey with non-expert participants, we researched intuitive unimodal gaze interaction concepts to complete a series of interaction tasks with a 3D UI in a HoloLens 2 emulation of a 3D holographic passenger display inside a car.

Keywords: gaze control, user elicitation, holographic interfaces, interaction design, user centered design, automotive HMI, 3D gaze interaction

Index Terms: Human-centered computing – Human computer interaction (HCI) – Interaction design – Interaction design process and methods – User centered design

1 INTRODUCTION

Current market reports expect a substantial increase in the so-called metaverse and Extended Reality (XR) or Cross Reality (CR) applications in the following years [42, 44]. Advancements in the field of autonomous driving technology also introduce new opportunities for immersive, non-driving-related infotainment and interaction experiences in future car interiors [48]. The demonstration of the head-worn CR device Microsoft HoloLens 2 in a moving vehicle shows a promising vision of how navigational content and HMI controls could be experienced in mixed reality while driving [4]. Augmented Reality HUDs (AR-HUDs), which have become state-of-the-art in multiple modern vehicles [39], provide an additional platform for CR experiences. Furthermore, the development of immersive 3D display technologies, including holographic 3D displays [12], can further expand the design space

---

* maryia.kazhura@volkswagen.de
* bea.vorhof@student.uni-luebeck.de
* andre.calerovaldez@uni-luebeck.de

for in-car interactions. In a holographic 3D display, virtual content is displayed in full natural 3D and placed in a large depth range in front of or behind a physical display screen, without loss of resolution or the need of extra head-worn hardware [38]. In addition to an increased interest in touchless interaction experiences in light of the recent COVID-19 pandemic [16, 33], these developments encourage research endeavors towards new interaction modes and design ideas for 3D CR interfaces.

Gaze is an inclusive modality that enables a broader range of potential users to interact with computer systems and CR applications. Gaze interaction can reduce the physical discomfort associated with prolonged and unsupported mid-air gestural input [14, 20] and allow users to interact with the system while using their hands for a different activity [37]. Since attention and eye movements go hand-in-hand [6, 15], explicit gaze interactions can benefit applications that require continuous attention [22], which would be particularly useful in autonomous driving scenarios or non-driving passengers. Gaze patterns can reveal specific areas of interest which can enable individualized UIs inside the vehicle [15]. Implicit gaze controls can be used to predict which object user intends to interact with, thus allowing the system to adjust the UI accordingly [13]. Some studies associated eye gaze interaction with faster task completion times compared to multimodal gaze+voice input [31], gaze input combined with a physical button or controller confirmation [35], hand pointing [43] or mouse selection [40]. Past studies also reported a lower cognitive workload of eye gaze compared to multimodal gaze+voice control [31] and a higher perceived ease compared to multimodal gaze+controller or gaze+button interaction [35], albeit with higher reported instances of eye fatigue and faster rising eye fatigue levels compared to head gaze [36]. Some researchers argue that the benefit of gaze control depends on the task at hand, reducing cognitive load in some tasks while increasing the load in others [11]. A general recurring challenge of applications that involve gaze controls is the so-called Midas Touch problem [17], which refers to users accidentally selecting items simply by looking at them [7].

In this contribution, we present and discuss unimodal gaze interaction proposals for eight exemplary interaction tasks with a potential holographic 3D display interface.

2 METHOD

This research project was carried out in three phases. In Phase 1, we reviewed literature to identify suitable gaze interaction methods. In Phase 2, we conducted an elicitation survey with non-expert participants to collect gaze interaction proposals for eight 3D interaction tasks. In Phase 3, we analyzed the results from the elicitation survey and binned the proposed interaction ideas. In Phase 4, we deduced the interaction concepts based on results from the elicitation survey, implemented each concept using the Microsoft MRTK [25, 26] version 2.7.2 and integrated the resulting UI with a HoloLens 2 emulation of a 240 × 135 mm large holographic 3D display inside a car.

2.1 Review of gaze interaction methods (Phase 1)

We identified nine different gaze interaction methods, which we considered suitable for use with our 3D display interface: (1) Eye pointing to trigger selections or display object-related information [2, 17, 23, 40]; (2) Smooth gaze pursuits of moving targets [9, 45]; (3) Gaze and blinking [24]; (4) Gaze and nodding [3]; (5) Combined head- and gaze-cursor movements [41]; (6) Gaze gestures [18, 28, 30], (7) Exploration and selection of menu items via sidebars for gaze selection [7]; (8) Dwell-Time techniques [8, 30, 32, 34, 47], for example to lock the degrees of freedom during 3D object manipulation [21], with dwell visualization presented on the object directly or around the cursor [10] and (9) locking of gaze cursor to a target area [29, 49].

2.2 Interaction tasks and priming

We selected eight interaction tasks based on a previous work which explored use cases, interaction tasks and suitable hand gestures for interaction with in-car holographic 3D displays [19]. The selected tasks were presented in the following order (1-8):

- *Search in menu:* Turn a menu wheel until a desired item is at the center.
- *Select item:* Select a desired item from a selection of multiple items.
- *Pull object closer:* Move a 3D object in depth in egocentric perspective towards self, until it reaches a desired position.
- *Rotate object:* Rotate a 3D object around its vertical axis until it reaches a desired rotation.
- *Move slider:* Move a temperature slider horizontally until it reaches a desired temperature.
- *Move map:* Move a 3D map in bird-view perspective until the map is centered around a target location.
- *Zoom map:* Zoom a 3D map in bird-view perspective until the desired view is reached.
- *Select a target on a map:* Select a desired target marker on the 3D map in bird-view perspective.

To appropriately prime participants to the affordances of an in-car holographic 3D display, we prepared extensive video explanations of the nine selected unimodal gaze interaction methods, as well as a step-by-step visual representation of each task execution. The task visualization included rendered images of the target 3D UI embedded in a hypothetical car interior displaying the following steps: default state of the UI prior to interaction, visual representation of the active target, visual representation of the selected target, visualization of the target as it is being moved/rotated/zoomed, visualization of the released but still active target when it is no longer being manipulated and the end of the interaction with the target being no longer activated.

2.3 Sample

We split the elicitation survey into two parts with four tasks each, to reduce the time demand and workload for each survey participant. Hence, we recruited two groups of volunteering participants who were compensated for their effort with gift vouchers: Group A (*N* = 34, F =15, M = 19) completed tasks one to four, Group B (*N* = 30, F =14, M = 16) completed tasks five to eight. Most participants (Group A: 88 %; Group B: 73%) used gaze controls less than once per year or never. The age distribution of participants in each group is depicted in Figure 2, with 13 participants in Group A and 13 participants in Group B being born between 1982 and 1964.

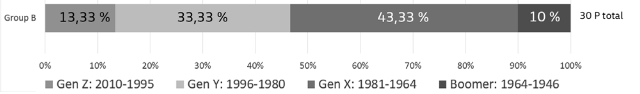

Figure 2: Age distribution of the sample in Group A and B.

2.4 Elicitation survey (Phase 2)

To collect the proposals, we conducted the survey remotely. Participants received a virtual presentation that contained information about the aims and background of the study, holographic 3D displays and how they could be used in a car, videos presenting the gaze interaction methods described above, and an introduction to each task's goal and its step-by-step visualization. Data was collected using online questionnaires linked in the presentation. The first questionnaire gathered information about participants' age, gender, driving habits, technology usage and media consumption. After filling out the demographic questionnaire, participants would proceed to view the material about the first of four tasks. A link to the task questionnaire was placed at the end of each task description. In the task questionnaire, participants were asked to describe their gaze interaction idea, using the following instruction: "How would you complete this task using only your eye-gaze to control the user interface?". Participants were encouraged to refer to the gaze interaction methods described in the presentation. In addition, participants were asked to rate the ease and self-descriptiveness of their proposed interaction method on a 10-point scale. Ease referred to how easy the interaction would be to execute, while self-descriptiveness was the degree to which the interaction is self-explanatory, as described in the standard ISO 9241-110. At the end of each task questionnaire, participants rated the usefulness of unimodal gaze interaction for the completion of the given task.

3 RESULTS

3.1 Analysis of elicitation survey results (Phase 3)

In a first step, the 179 proposed interaction ideas were reviewed and filtered to remove incomplete responses that failed to describe a gaze interaction. The resulting number of analyzed proposals for each task was as follows: *Search in menu* (22), *Select item* (17), *Pull object closer* (14), *Rotate object* (15), *Move slider* (20), *Move map* (15), *Zoom map* (13) and *Select a target on a map* (13). The remaining responses were in part extensive and detailed, allowing us to use the set of responses to deduce meaningful interaction concepts and providing a valuable insight into participants' mental models. The number of gaze usefulness ratings obtained from each task was: *Search in menu* (23), *Select item* (21), *Pull object closer* (17), *Rotate object* (19), *Move slider* (24), *Move map* (17), *Zoom map* (17) and *Select a target on a map* (17). This shows how only a portion of the recruited participants completed all online questionnaires and provided usable data.

We binned the filtered proposals based on their similarity, so that proposals that described similar methods used during the single stages of interaction were binned into the same group. For example: we grouped proposals that described looking at various additional UI objects or buttons to move or rotate an element (3D arrows, a 3D scale, buttons with icons, etc.) into the same category. We then ranked the binned groups according to their average ratings of ease and self-descriptiveness. If multiple categories had an equal number of occurrences, we prioritized the category with a higher ease and self-descriptiveness rating.

With respect to the obtained gaze usefulness ratings (Figure 3), we found that gaze control was considered most useful for selection tasks, while being rated less useful for translation tasks like *Zoom Map*, *Search in Menu*, *Rotate Object* and P*ull Object closer*.

3.2 Final interaction concepts (Phase 4)

We deducted the final concepts presented in Figure 1 based on the top three categories from the binning procedure, but considering additional factor such as technical feasibility, ergonomics, and affordances of holographic 3D-displays. We then implemented the interaction concepts according to the description below, using the Microsoft MRTK.

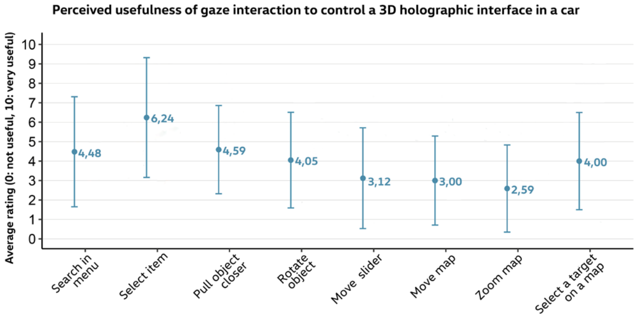

Figure 3: Average gaze usefulness ratings across eight tasks. Error bars depict standard deviations.

As a general rule of each concept design, we defined that the users' gaze point must be inside the UI area within the bounds of the virtual display, to enable any interaction with the system. Based on literature review and pilot tests with 10 participants, we selected a threshold of 300 ms for gaze fixation that triggered visual highlighting of the fixated object and the start of a dwell timer for another 1000 ms. Gaze selections therefore required a total fixation duration of 1300 ms. Audio feedback was additionally used to confirm most selections. All visual feedback for the gaze interactions was designed according to insights from literature (e.g. [10]) and recommendations from the MRTK [27]. Looking away from the virtual display resulted in a deactivation of all selected objects. Our final interaction concepts for each task are:

- *Search in menu:* User looks on the arrow above the menu to select the direction in which the wheel is rotated. Upon gaze contact, the arrow outlines are highlighted. At fixation, the arrow gradually turns green until for the duration of the dwell until selection is confirmed. User can then look at the selected arrow to rotate the menu wheel for the duration of arrow fixation. The arrow remains selected, until the gaze leaves the UI area, or another arrow is selected.
- *Select item:* Upon gaze contact, the item outlines are highlighted. At fixation, the button gradually turns green for the duration of the dwell, until selection is confirmed.
- *Pull object closer:* Upon gaze contact, the outlines of the box collider around the 3D object are highlighted. At fixation, the dwell timer initiates with the visualization of a wheel around the cursor that gradually fills until selection is confirmed. After selection, a 3D scale appears below the object, with a nearer arrow at the front and a farther arrow behind the object. When user fixates an arrow, it gradually turns green for the duration of the dwell with additional visualization of a wheel that gradually fills around the cursor. User can look at the selected arrow to move the object in the arrow direction for the duration of gaze fixation on the arrow. The arrow remains selected, until the gaze leaves the UI area within the display, or another arrow is selected.
- *Rotate object:* Upon gaze contact, the outlines of the box collider around the 3D object are highlighted and an extra 3D rotation icon is displayed in the lower left corner of the box. At fixation on the rotation icon, the dwell timer initiates with the visualization of a wheel around the cursor that gradually fills until selection is confirmed. Upon selection, the rotation icon disappears, and two rotation arrows appear below the box (one for clockwise and one for counterclockwise rotation around the vertical axis). Upon fixation of one arrow, it turns green, and user

can rotate the object in the arrow direction for the duration of arrow fixation.

- *Move slider:* Upon gaze contact, the outlines of the slider box are highlighted. At fixation, the box gradually turns green for the duration of the dwell, until selection is confirmed. User then fixates a point on the slider bar to trigger another dwell timer with the wheel visualization, until the dwell is completed, and the slider gradually moves to the selected position. User must select the slider box again to initiate a new movement.
- *Move map:* Upon gaze fixation of the map, four bars appear along the map edges. At fixation of one bar, it gradually turns green for the duration of the dwell. After the selection is confirmed, user can move the map by fixating the selected edge. The selected bar remains selected until the gaze leaves the UI area within the display, or another object is selected. Once the target location is in view, the user can select it via dwell and the map automatically centers around the selected location.
- *Zoom map (ZM):* Upon gaze fixation of the map, user can look at one of the zoom buttons near the right edge of the map. Upon gaze contact, the outlines of the zoom button are highlighted and at fixation, the button gradually turns green for the duration of the dwell until selection is confirmed. User can zoom in or out (depending on the selected button) by fixating the selected button. The button remains selected, until user's gaze leaves the map or another button is selected.
- *Select a target on a map:* Upon gaze contact, the outlines of the marker are highlighted. At fixation, the marker gradually turns green for the duration of the dwell, until selection is confirmed.

4 DISCUSSION

Our interaction concepts are deduced from the proposed ideas of non-expert participants, based on affordance guided priming material. Our step-by-step visualization of each task left it open to interpretation, as to whether the task could be completed in a continuous or discrete manner. Some participants proposed discrete methods for tasks that others would resolve continuously (e.g., moving an object in steps vs. moving it continuously). We might have gotten different results had we used other visuals to convey the task goals. While we expect our interaction concepts to be intuitive, self-descriptive, and easy to use based on the ratings of the elicited proposals, we suggest viewing the results from such surveys as inspiration and user centered guidance rather than explicit design instructions.

4.1 Future work

Since gaze can be efficiently combined with gesture or voice input to solve specific challenges in AR/VR applications [37, 46], our survey further collected interaction proposals for potential multimodal controls of the suggested tasks, combining gaze input with mid-air gestures, voice control, or other modalities that participants could imagine using in combination with gaze. Using the same approach, we deduced multimodal gaze-supported interaction concepts with mid-air gestures and voice commands. While these results are out of scope for this contribution, we plan to evaluate the task performance, user experience, and usability of the presented unimodal gaze interactions and compare them with a multimodal gaze + hand tracking interface. Furthermore, we will compare hand tracking-based interactions described in Kazhura ([19]) with multimodal gaze-hand and gaze-voice interactions to further investigate the potential benefits of gaze input for interaction with immersive 3D UIs.

4.2 Limitations

One major limitation of our approach is the remote elicitation approach, which required participants to elicit ideas for a technology they had little to no experience with, based on low-fidelity priming materials (descriptions, videos, and images as opposed to an immersive CR experience of each task's affordances). This low-fidelity approach could have limited participants' creativity and the perception of gaze controls usefulness for 3D UIs. Furthermore, the generalizability of our results is limited by the specific priming material used to elicit the proposals, as the proposals were heavily influenced by the affordances of the visualized UI. In addition, the study was anonymous, thus limiting communication between us and participants. Participants could not make quick inquiries about the presented methods or openly discuss a specific topic or idea. We can see how this impacted our results based on the number of complete and usable responses, in contrast to the total sample size of each group. It is also important to note that while we surveyed non-experts and split the tasks into two groups, experienced AR/VR users might propose different and more consistent approaches with a greater focus on established UI/UX guidelines and technical feasibility. The cultural diversity of our sample was also limited, since we recruited participants from a single region in Germany. People with other demographic backgrounds may suggest different interaction ideas or dismiss certain ideas entirely. The gaze method 9 (locking of gaze cursor to a target area) was presented in less detail compared to the other methods, which may have made it more challenging for participants to understand and include the method in their proposals. Finally, the presented UIs don't include much visual clutter. It is therefore debatable whether the concepts are transferable to more complex holographic interfaces.

5 CONCLUSION

Non-experts who have had little experience with XR or CR technologies may not see the benefits of gaze control for certain tasks yet and may find it challenging to propose suitable interaction ideas. We found the results from the elicitation study very helpful in preparing user-centered gaze control-based 3D UIs for future evaluation. Based on our experience, we encourage the inclusion of participants' ideas in early design iterations of CR-UIs. However, we recommend considering a more direct and interactive CR survey approach, to help participants whose imagination and creativity may be limited. For example, by letting users experience the priming UI in a CR application in real-time using immersive authoring tools or by using a visual editor on a tablet. In any case, it is important to carefully prime non-experts with respect to possible approaches to obtain more meaningful results.

ACKNOWLEDGMENTS

This study was funded by the Volkswagen Aktiengesellschaft. The results, opinions and conclusions expressed in this publication are not necessarily those of Volkswagen Aktiengesellschaft.

REFERENCES

[1] Adam, T. and Untaroiu, C. D. 2011. Identification of occupant posture using a Bayesian classification methodology to reduce the risk of injury in a collision. *Transportation Research Part C: Emerging Technologies* 19, 6, 1078–1094.

[2] Alt, F., Schneegass, S., Auda, J., Rzayev, R., and Broy, N. 2014. Using Eye-Tracking to Support Interaction with Layered 3D Interfaces on Stereoscopic Displays. In *Proceedings of the 19th International Conference on Intelligent User Interfaces*. IUI '14. Association for Computing Machinery, New York, NY, USA, 267–272. DOI=10.1145/2557500.2557518.

[3] Ataya, A., Kim, W., Elsharkawy, A., and Kim, S. Gaze-Head Input: Examining Potential Interaction with Immediate Experience Sampling in an Autonomous Vehicle. *Applied Sciences* 10, 24, 9011.

[4] Bach, D. 2022. *With their HoloLens 2 project, Microsoft and Volkswagen collaborate to put augmented reality glasses in motion*.

[5] Bose, D., Crandall, J. R., Untaroiu, C. D., and Maslen, E. H. 2010. Influence of pre-collision occupant parameters on injury outcome in a frontal collision. *Accident; analysis and prevention* 42, 4, 1398–1407.

[6] Duchowski, A. T. 2017. *Eye tracking methodology. Theory and practice*. Springer, Cham.

[7] Elmadjian, C. and Morimoto, C. H. 2021. GazeBar: Exploiting the Midas Touch in Gaze Interaction. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI EA '21. Association for Computing Machinery, New York, NY, USA. DOI=10.1145/3411763.3451703.

[8] Esteves, A., Shin, Y., and Oakley, I. Comparing selection mechanisms for gaze input techniques in head-mounted displays. *International Journal of Human-Computer Studies* 139, 102414.

[9] Esteves, A., Verweij, D., Suraiya, L., Islam, R., Lee, Y., and Oakley, I. 2017. SmoothMoves: Smooth Pursuits Head Movements for Augmented Reality. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. UIST '17. Association for Computing Machinery, New York, NY, USA, 167–178. DOI=10.1145/3126594.3126616.

[10] Fernandez, M., Mathis, F., and Khamis, M. 2020. GazeWheels: Comparing Dwell-Time Feedback and Methods for Gaze Input. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*. NordiCHI '20. Association for Computing Machinery, New York, NY, USA. DOI=10.1145/3419249.3420122.

[11] Gauselmann, P., Frings, C., and Tempel, T. 2023. Cognitive offloading benefits eye gaze interaction. *Applied Cognitive Psychology*, 1–5.

[12] Häussler, R., Gritsai, Y., Zschau, E., Missbach, R., Sahm, H., Stock, M., and Stolle, H. Large real-time holographic 3D displays: enabling components and results. *Applied optics* 56, 13, F45-F52.

[13] Henderson, J. M. 2017. Gaze Control as Prediction. *Trends in cognitive sciences* 21, 1, 15–23.

[14] Hincapié-Ramos, J. D., Guo, X., Moghadasian, P., and Irani, P. 2014. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *CHI 2014, one of a CHInd. Conference proceedings : Toronto, Canada, April 26 - May 1, 2014 ; the 32nd Annual ACM Conference on Human Factors in Computing Systems*. Assoc. for Computing Machinery, New York, NY, 1063–1072. DOI=10.1145/2556288.2557130.

[15] Horrey, W. J., Wickens, C. D., and Consalus, K. P. 2006. Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. *Journal of experimental psychology. Applied* 12, 2, 67–78.

[16] Huang, S., Ranganathan, S. P. B., and Parsons, I. 2020. To Touch or Not to Touch? Comparing Touch, Mid-Air Gesture, Mid-Air Haptics for Public Display in Post COVID-19 Society. In *SIGGRAPH Asia 2020 Posters*. SA '20. Association for Computing Machinery, New York, NY, USA. DOI=10.1145/3415264.3425438.

[17] Jacob, R. J. K. 1990. What You Look at is What You Get: Eye Movement-Based Interaction Techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '90. Association for Computing Machinery, New York, NY, USA, 11–18. DOI=10.1145/97243.97246.

[18] Jenke, M., Huppenbauer, L., and Maier, T. 2018. Investigation of continual operation procedures via a user centered gaze control by means of flexible gaze gestures. *Zeitschrift für Arbeitswissenschaft* 72, 1, 23–34.

[19] Kazhura, M. 2022. Exploring New Depths: How Could Passengers Interact with Future In-Car Holographic 3D Displays? In *HCI in Mobility, Transport, and Automotive Systems: 4th International Conference, MobiTAS 2022, Held as Part of the 24th HCI International Conference, HCII 2022, Virtual Event, June 26 - July 1, 2022, Proceedings*. Springer-Verlag, Berlin, Heidelberg, 35–61. DOI=10.1007/978-3-031-04987-3.

[20] Kim, J. H., Ari, H., Madasu, C., and Hwang, J. 2020. Evaluation of Hologram Distances in Reducing Shoulder Stress During Augmented Reality Interactions. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 64, 1, 868–871.

[21] Liu, C., Plopski, A., and Orlosky, J. OrthoGaze: Gaze-based three-dimensional object manipulation using orthogonal planes. *Computers & Graphics* 89, 1–10.

[22] Lu, F., Davari, S., Lisle, L., Li, Y., and Bowman, D. A. 2020. Glanceable AR: Evaluating Information Access Methods for Head-Worn Augmented Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces. VR 2020 : proceedings : 22-26 March 2020, Atlanta, Georgia*. IEEE, Piscataway, NJ, 930–939. DOI=10.1109/VR46266.2020.00113.

[23] McNamara, A. and Kabeerdoss, C. 2016. Mobile Augmented Reality: Placing Labels Based on Gaze Position. In *Adjunct proceedings of the 2016 IEEE International Symposium on Mixed and Augmented Reality. Proceedings*. IEEE, Piscataway, NJ, 36–37. DOI=10.1109/ISMAR-Adjunct.2016.0033.

[24] Meng, X., Xu, W., and Liang, H.-N. 2022. An Exploration of Hands-free Text Selection for Virtual Reality Head-Mounted Displays. In *2022 IEEE International Symposium on Mixed and Augmented Reality. 17-21 October 2022, Singapore*. IEEE, Piscataway, NJ, 74–81. DOI=10.1109/ISMAR55827.2022.00021.

[25] Microsoft. 2021. *Mixed Reality Toolkit*. Unity Software Package, Redmont, USA.

[26] Microsoft. 2022. *What is Mixed Reality Toolkit 2?* Online article. https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/?view=mrtkunity-2022-05. Accessed 2 March 2023.

[27] Microsoft. 2023. *Eye-gaze-based interaction on HoloLens 2*. Online Article. https://learn.microsoft.com/en-us/windows/mixed-reality/design/eye-gaze-interaction. Accessed 8 July 2023.

[28] Mllenbach, E., Lillholm, M., Gail, A., and Hansen, J. P. 2010. Single Gaze Gestures. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*. ETRA '10. Association for Computing Machinery, New York, NY, USA, 177–180. DOI=10.1145/1743666.1743710.

[29] Murata, A. and Karwowski, W. Automatic Lock of Cursor Movement: Implications for an Efficient Eye-Gaze Input Method for Drag and Menu Selection. *IEEE Trans. Human-Mach. Syst.* 49, 3, 259–267.

[30] Namnakani, O., Abdrabou, Y., Grizou, J., Esteves, A., and Khamis, M. 2023. Comparing Dwell Time, Pursuits and Gaze Gestures for Gaze Interaction on Handheld Mobile Devices. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Association for Computing Machinery, New York, NY, USA. DOI=10.1145/3544548.3580871.

[31] Parisay, M., Poullis, C., and Kersten-Oertel, M. 2021. EyeTAP: Introducing a multimodal gaze-based technique using voice inputs with a comparative analysis of selection techniques. *International Journal of Human-Computer Studies* 154, 102676.

[32] Paulus, Y. T. and Remijn, G. B. Usability of various dwell times for eye-gaze-based object selection with eye tracking. *Displays* 67, 101997.

[33] Pearson, J., Bailey, G., Robinson, S., Jones, M., Owen, T., Zhang, C., Reitmaier, T., Steer, C., Carter, A., Sahoo, D. R., and Raju, D. K. 2022. Can't Touch This: Rethinking Public Technology in a COVID-19 Era. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. Association for Computing Machinery, New York, NY, USA. DOI=10.1145/3491102.3501980.

[34] Pfeuffer, K., Abdrabou, Y., Esteves, A., Rivu, R., Abdelrahman, Y., Meitner, S., Saadi, A., and Alt, F. 2021. ARtention. A design space for gaze-adaptive user interfaces in augmented reality. *Computers & Graphics* 95, 4, 1–12.

[35] Pfeuffer, K., Mecke, L., Delgado Rodriguez, S., Hassib, M., Maier, H., and Alt, F. 2020. Empirical Evaluation of Gaze-Enhanced Menus in Virtual Reality. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*. VRST '20. Association for

Computing Machinery, New York, NY, USA. DOI=10.1145/3385956.3418962.

[36] Qian, Y. Y. and Teather, R. J. 2017. The Eyes Don't Have It: An Empirical Comparison of Head-Based and Eye-Based Selection in Virtual Reality. In *Proceedings of the 5th Symposium on Spatial User Interaction*. SUI '17. Association for Computing Machinery, New York, NY, USA, 91–98. DOI=10.1145/3131277.3132182.

[37] Rakkolainen, I., Farooq, A., Kangas, J., Hakulinen, J., Rantala, J., Turunen, M., and Raisamo, R. 2021. Technologies for Multimodal Interaction in Extended Reality—A Scoping Review. *MTI* 5, 12, 81.

[38] Reichelt, S., Häussler, R., Fütterer, G., and Leister, N. 2010. Depth cues in human visual perception and their realization in 3D displays. In *Proceedings Volume 7690: Three-Dimensional Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics IV*. SPIE. DOI=10.1117/12.850094.

[39] Reportlinker.com. 2022. *Automotive Head-up Display (HUD) Industry Report, 2022* 6318548.

[40] Sibert, L. E. and Jacob, R. J. K. 2000. Evaluation of Eye Gaze Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '00. Association for Computing Machinery, New York, NY, USA, 281–288. DOI=10.1145/332040.332445.

[41] Sidenmark, L., Potts, D., Bapisch, B., and Gellersen, H. 2021. Radi-Eye: Hands-Free Radial Interfaces for 3D Interaction Using Gaze-Activated Head-Crossing. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. Association for Computing Machinery, New York, NY, USA. DOI=10.1145/3411764.3445697.

[42] Statista, Inc. 2022. *AR Hardware - Worldwide*. Online article. https://www.statista.com/outlook/amo/ar-vr/ar-hardware/worldwide.

[43] Tanriverdi, V. and Jacob, R. J. K. 2000. Interacting with Eye Movements in Virtual Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '00. Association for Computing Machinery, New York, NY, USA, 265–272. DOI=10.1145/332040.332443.

[44] Trenker, J. 2023, February. *Metaverse Market Report*. [Report]. Statista.

[45] Vidal, M., Bulling, A., and Gellersen, H. 2013. Pursuits: Spontaneous Interaction with Displays Based on Smooth Pursuit Eye Movement and Moving Targets. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UbiComp '13. Association for Computing Machinery, New York, NY, USA, 439–448. DOI=10.1145/2493432.2493477.

[46] Wang, Z., Wang, H., Yu, H., and Lu, F. Interaction With Gaze, Gesture, and Speech in a Flexibly Configurable Augmented Reality System. *IEEE Trans. Human-Mach. Syst.* 51, 5, 524–534.

[47] Ware, C. and Mikaelian, H. H. An evaluation of an eye tracker as a device for computer input2. *SIGCHI Bull.* 17, SI, 183–188.

[48] Wiegand, G., Mai, C., Holländer, K., and Hussmann, H. 2019. InCarAR: A Design Space Towards 3D Augmented Reality Applications in Vehicles. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. AutomotiveUI '19. Association for Computing Machinery, New York, NY, USA, 1–13. DOI=10.1145/3342197.3344539.

[49] Zhang, S., Tian, Y., Wang, C., and Wei, K. Target selection by gaze pointing and manual confirmation: performance improved by locking the gaze cursor. *Ergonomics* 63, 7, 884–895.