



OPEN ACCESS

EDITED BY

Antonio Sarasa-Cabezuelo,
Complutense University of Madrid, Spain

REVIEWED BY

Gavindya Jayawardena,
The University of Texas at Austin, United States
Xinjie Wang,
National University of Defense Technology,
China

*CORRESPONDENCE

Björn R. Severitt,
✉ bjoern.severitt@uni-tuebingen.de

RECEIVED 14 February 2025

ACCEPTED 26 March 2025

PUBLISHED 16 April 2025

CITATION

Severitt BR, Sauer Y, Neugebauer A, Agarwala R,
Castner N and Wahl S (2025) The interplay of
user preference and precision in different gaze-
based interaction methods in
virtual environments.

Front. Virtual Real. 6:1576962.

doi: 10.3389/frvir.2025.1576962

COPYRIGHT

© 2025 Severitt, Sauer, Neugebauer, Agarwala,
Castner and Wahl. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

The interplay of user preference and precision in different gaze-based interaction methods in virtual environments

Björn R. Severitt^{1*}, Yannick Sauer², Alexander Neugebauer¹,
Rajat Agarwala¹, Nora Castner² and Siegfried Wahl^{1,2}

¹ZEISS Vision Science Lab, University of Tübingen, Tübingen, Germany, ²Carl Zeiss Vision International GmbH, Aalen, Germany

Introduction: Extended reality (XR) technologies, particularly gaze-based interaction methods, have evolved significantly in recent years to improve accessibility and reach broader user communities. While previous research has improved the simplicity and inclusivity of gaze-based choice, the adaptability of such systems – particularly in terms of user comfort and fault tolerance – has not yet been fully explored.

Methods: In this study, four gaze-based interaction techniques were examined in a visual search game in virtual reality (VR). A total of 52 participants were involved. The techniques tested included selection by dwell time, confirmation by head orientation, nodding and smooth pursuit eye movements. Both subjective and objective performance measures were assessed, using the NASA-TLX for perceived task load and time to complete the task and score for objective evaluation.

Results: Significant differences were found between the interaction techniques in terms of NASA-TLX dimensions, target search time and overall performance. The results indicate different levels of efficiency and intuitiveness of each method. Gender differences in interaction preferences and cognitive load were also found.

Discussion: These findings highlight the importance of personalizing gaze-based VR interfaces to the individual user to improve accessibility, reduce cognitive load and enhance the user experience. Personalizing gaze interaction strategies can support more inclusive and effective VR systems that benefit both general and accessibility-focused populations.

KEYWORDS

eye tracking, gaze-based interaction, gaze, communication, accessibility, virtual reality

1 Introduction

Humans naturally rely on their gaze to perceive, explore, and interact with their environment. Extending this ability to system interaction via gaze as an input method seems intuitive, as it reflects our natural focus on objects or regions of interest. However, developing a system that correctly interprets gaze still presents major challenges. Systems must not only accurately capture gaze, but also interpret the user's intentions in the correct interaction syntax. There are already a number of systems that have effectively incorporated gaze input for support in immersive games (Smith and Graham, 2006), smart

wearables (Mastrangelo et al., 2018), and even medical devices that support communication or mobility (Pannasch et al., 2008; Subramanian et al., 2019). However, gaze can still be pushed aside in favor of the more standard (e.g., touch or physical buttons) or other communication approaches in human-machine interfaces. These interface design choices are applicable for any number of systems, but can also limit them to certain users. For instance, individuals with no motor disabilities or who have their hands available at any given moment. Therefore, given the challenges gaze input presents, it is still one of the more implicit modalities for hands-free system interaction.

Often, communication between user and system comes with customization challenges that can affect how accessible the experience can be (Dey et al., 2019; Macaranas et al., 2015). Traditional communication approaches can be bound to physical actions, w.r.t. button clicks or scrolling with a mouse, touch gestures like swiping or pinching. Or more sophisticated approaches that couple gaze to a gesture can be more natural and reduce strain (Pfeuffer et al., 2017b). But the pitfalls of these communication approaches are that they are exclusive to able-bodied individuals (Gür et al., 2020). Verbal-based modalities are an ever-increasing alternative, especially with the recent boom in Natural Language Processing (NLP) models. This approach can quickly extract the context in which the user wishes to accomplish their specific task: The best example being chatbots. Despite these advantages, NLP-based systems face other challenges. Mainly, there are sometimes unpredictable outputs from the *black box*-style of how the model learns (Lin et al., 2023; Gou et al., 2023). For a user, this can result in a system being slightly annoying, but also safety critical. For example, controlling a robotic arm can be susceptible to a user not appropriately describing the object they wish to have the robot grasp. They may not forget to specify the tea kettle should be grabbed by the handle since they plan to pour it into a cup. The challenge of understanding a user's implicit or unconscious understanding of a given context is precisely where gaze input can supplement User Interfaces (UI's). Going back to the tea kettle example, a user would fixate momentarily on the handle moments before proceeding to other look ahead fixations indicating the next steps in the task (Bovo et al., 2020; Pelz and Canosa, 2001). This fixation input can be enough to communicate to a system, *grasp here*.

Though eye tracking technology for gaze-based communication has been around for nearly 40 years, it is still a modality that is highly accessible and natural. It can be performed with relatively low-cost devices that offer high accuracy (within 1°) (Lee et al., 2020; Rakhmatulin, 2020) and can provide scene analysis through efficient object detection methods (Jha et al., 2021). Eye tracking integrated into Virtual Reality (VR) brings this technology to a new level of users, as VR headsets are more commercially available. However, as with any input modality, there are challenges; mainly how to distinguish between an accidental gaze and a deliberate command—*The Midas Touch Problem*. This issue has prompted numerous solutions, which we will overview in Section 2, but one aspect that is underexplored in the previous literature is more detailed aspects of the user experience behind these approaches, such as engagement, threshold preference, and tolerance to error.

Considering that a satisfactory user experience is important for a user's opinion of whether an extended reality is useful, comfortable, or intuitive, we wanted to investigate aspects of different gaze-based

interaction paradigms could contribute to the overall experience. In this paper, we present and evaluate four common gaze- and head-movement-based interaction methods and compare their performance in an immersive environment where participants perform a gamified visual search task. Specifically, we investigated gaze-based selection by dwell time and methods in which gaze-based selection is confirmed by head direction, nodding, or a specific eye movement. By studying these paradigms in terms of both objective and subjective factors, such as task performance, signal accuracy, and preference, we aim to identify interaction strategies that maximize the efficiency and usability of gaze-driven systems and promote customization and adaptability for a range of users. Understanding user preferences and precision in gaze-based selection is important to developing intuitive and efficient interaction methods in VR. While gaze provides a natural, hands-free input modality, differences in accuracy, comfort, and cognitive load on users can affect usability. By evaluating both subjective preferences and objective precision, we can identify strategies that reduce selection errors, increase user satisfaction and improve accessibility. This ensures that gaze-based interfaces are not only technically effective, but also adaptable to different user needs, ultimately leading to a more inclusive and user-friendly VR experience.

2 Related work

At a foundational level, gaze as an input for various interface tasks can be simple and effective, such as for text entry (Hansen et al., 2004; Majaranta and Bates, 2009; Majaranta and Riih a, 2002; Ward and MacKay, 2002; Wobbrock et al., 2008) and PIN entry (Best and Duchowski, 2016; Hoanca and Mock, 2006). It can also prove as a useful means of navigation, for instance with hierarchical interfaces (Huckauf and Urbina, 2007; Huckauf and Urbina, 2008) or predict the future position (Bremer et al., 2021). As this current research is focused on different methods of gaze-based selection, namely, in extended reality environments, we contain the rest of the related work to this aspect of gaze-based interaction. However, for further, more detailed reviews on more broader topics in gaze-based interaction, we refer readers to (Duchowski, 2018; Plopski et al., 2022).

The most intuitive way to select an object with gaze alone is the dwell time introduced by Jacob (1990). This can also be used as a metric for interest in an object (Starker and Bolt, 1990). As mentioned above, this method leads to the problem of the *Midas Touch*, as there is no way to recognize which look is a real look of interest. To address this challenge, additional modalities such as gestures have been explored, as demonstrated by Špakov and Majaranta (2012). They investigated the use of various head gestures for different activities, such as item selection and navigation, and found that users' preferences for gestures varied depending on the task. For selection, a nod was generally preferred, while head turning was favored for navigation, and tilting the head was most effective for switching functional modes. Another approach involves using specific eye movements for confirmation, first introduced by Vidal et al. (2013), who employed smooth pursuit eye movements—slow, continuous motions that allow the eyes to track moving objects—to select an

object. Esteves et al. (2015) later built on this technique, creating a spherical object that users could follow with their eyes to confirm their selection.

In VR, head gaze and eye gaze are often compared for their effectiveness (Pfeuffer et al., 2017a; Piumsomboon et al., 2017), yielding mixed results. Qian and Teather (2017) found that head gaze performed better, while Blattgerste et al. (2018) reached the opposite conclusion. They attributed their findings to the higher accuracy of the eye-tracking data used in their study, which made it easier for participants to interact with the system. Fernandes et al. (2025) investigated this comparison further by evaluating gaze, head gaze and controller input with different feedback methods. In this experiment, the participant must select targets by selecting them using the input method and confirming this by pressing a button on the controller. Their results suggest that gaze-based selection, when combined with appropriate feedback and a button press, can perform as well or better than controllers in certain Augmented Reality (AR) or VR tasks.

A logical next step is to combine both eye and head gaze methods, as explored by Sidenmark and Gellersen (2019). They proposed three different approaches for selecting a target by looking at it and confirming this selection by adjusting the head direction, e.g., by turning the head towards the object of interest. This combination enhances control and flexibility in the selection process. Wei et al. (2023) took a different approach to the use of eye and head gaze. They created a probabilistic model based on the endpoints of the gaze and used this to decide whether and which object should be selected.

Although previous studies have provided valuable insights into gaze-based and multimodal interaction techniques, there are still major gaps. Most previous research has focused on evaluating a single selection method, rather than comparing multiple approaches within the same experimental setting. In addition, many studies have been conducted in highly controlled environments with fixed spatial arrangements that do not reflect scenarios in which users have to repeatedly search for targets in different locations. Our study, while still conducted in a controlled environment, introduces a more dynamic task structure in an immersive VR environment that requires participants to efficiently locate targets instead of interacting with static elements. Furthermore, users' preferences and precision in gaze-based selection have not yet been sufficiently explored, especially in the context of immersive VR experiences. To address these limitations, our study introduces a comparative analysis of multiple gaze-based selection methods in an interactive VR environment. By evaluating both objective (performance, accuracy) and subjective (user experience, workload) factors, we aim to provide a more comprehensive understanding of how different gaze-based techniques perform for different users. Our findings will contribute to the development of more adaptable and user-friendly gaze interaction systems that better suit individual needs and preferences in VR.

3 Methods

To compare the gaze-based selection methods, we developed a custom VR game using Unity (Haas, 2014), integrated into the

VisionaryVR framework (Hosp et al., 2024). This setup allowed participants to test each method and then complete the NASA TLX questionnaire (Hart and Staveland, 1988) directly afterward. Our game was designed in such a way that there is a variable search task that the participants must find the target as quickly as possible in order to maximize their points (see Section 3.2). They play the game with each gaze-selection method in semi-randomized order. The procedure for each method began with an introduction, where the method and its adjustable parameters are explained. Then came the test phase, during which participants could try the method and adjust the parameters to their comfort. Once participants felt they had chosen their preferred parameters, the main phase could start. In the main phase, participants played the game for ten rounds. The data collected during this phase was later used for analysis. After the main phase, participants were shown the questionnaire scene to answer the NASA-TLX questions. This process was repeated for each method, with the gaze-dwell method always being the first to allow the participant to learn the game, and the other methods in random order.

3.1 Interaction methods

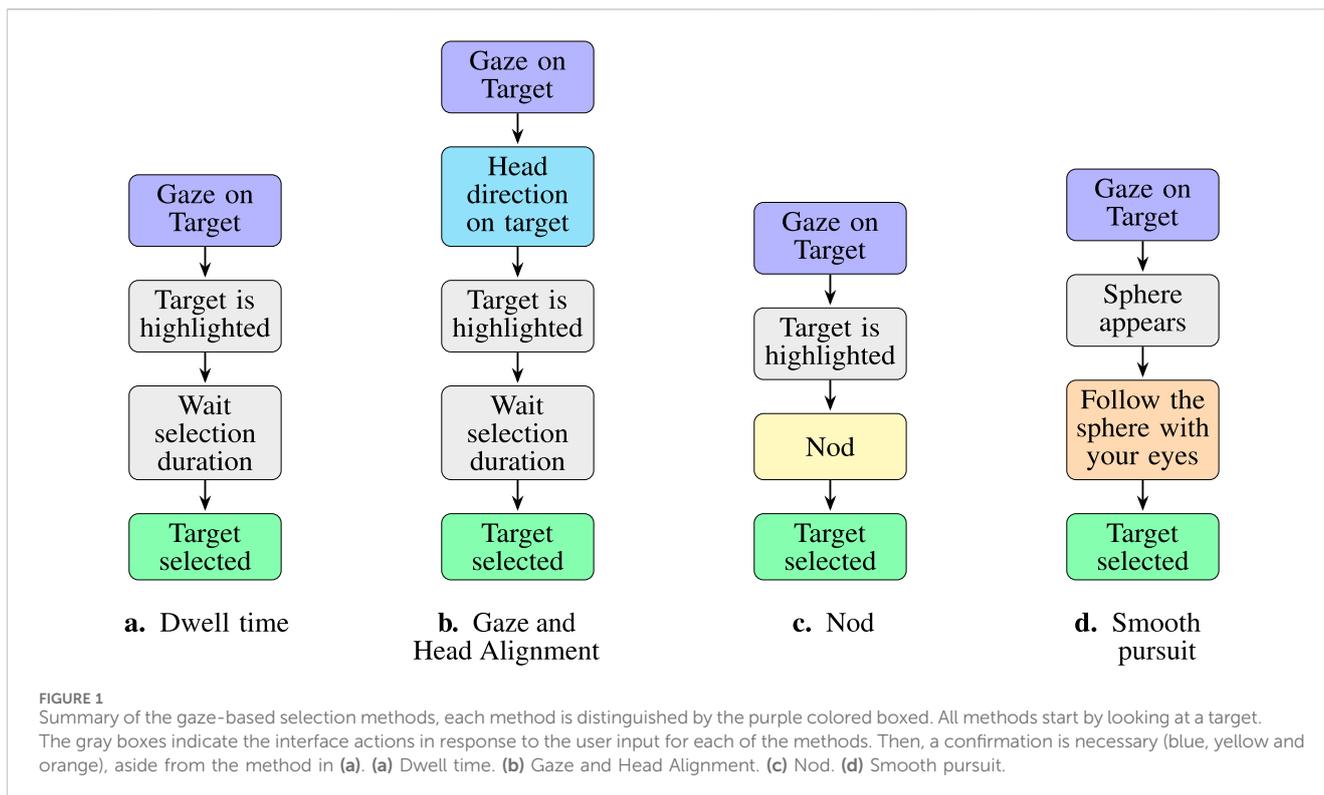
Each selection method uses gaze as the core selection technique, though we investigated differing modalities to confirm a user's selection. Below, we describe each method in detail and how participants could customize them to their preferences (see the additional Figure 1 depicting each method).

3.1.1 Gaze dwell

This method is the most common and is based exclusively on the dwell time introduced by Jacob (1990). Here, a target is selected by fixating on it for a sufficient duration. As the player gazes at an object, the target is highlighted with an outline. We implement gaze dwell as follows. The time spent looking at the target is visually represented by a change in the outline color, transitioning from green to red as the selection is locked in. For this method, the only parameter the participant must set is the *selection duration*, which specifies the time in seconds the user must gaze at the target for it to be selected. We set a minimum duration to 0.05 and a maximum duration to one and the step increments were 0.05.

3.1.2 Gaze and head

This method is inspired by the method from Sidenmark and Gellersen (2019). Here, the player must not only look at the target, but also confirm the selection by aligning their head with the target. The direction of the head was indicated by a green dot. The target is highlighted with a magenta outline when gaze is directed towards it, and with a green outline when head and gaze are directed to it. The green border changes to red to visualize the dwell time. For this method, the participant must set two parameters: (1) *selection duration* is similar as for the *dwell time* method, but now it represents the duration the gaze and the head direction have to be aligned. (2) *Head Orientation Precision* specifies the allowed offset in degree of the user's head orientation relative to the target for the selection confirmation. We set a minimum degree to two and a maximum degree to 20 and the step increments were 2. Thus, a higher degree indicates less precision the system needs to confirm selection.



3.1.3 Nod

This method is based on the results of Špakov and Majaranta (2012). The nodding method involves selecting the target by first gazing at it and then performing a nod gesture. The nod is recognized as a confirmation signal to complete the selection. For this method, the participant must set two parameters when the gesture is recognized as a nod: (1) *nod strength*, which specifies the amount of head movement in degree required for the system to register the nod gesture. (2) *Nod direction precision*, which specifies the allowed offset in degree of the final head position required to confirm the nod. For (1), we set a minimum degree to five and a maximum degree to 30 and the step increments were 1. For (2), we set a minimum degree to one and a maximum degree to 20 and the step increments were 1.

3.1.4 Smooth pursuit

This method follows the idea of Esteves et al. (2015). For our implementation, an orange-colored sphere appears after looking at the target and starts to move along a pre-defined path. The player must follow the sphere's movement with their gaze. If the player tracks the sphere and the smooth pursuit's trajectory and velocity align to the moving sphere, the target will be selected. This method has two different parameters to set: (1) *Tracking precision*, which represents the required correlation between the user's viewing direction and the movement of the sphere. We have set a minimum of 0.05 and a maximum of 1, with an increment of 0.05. (2) *Movement pattern*, which is the path that the sphere follows. Here the participant can choose between three options: *Circle*, *Bounding* and *Random Walk*.

3.2 Game

We evaluated each interaction method in the VR game environment that we developed. The goal was to get a high score. To accomplish this, they have to tweak the parameters for each interaction method to give them the best performance. Thus, they have to figure out factors such as comfort, speed, and accuracy that can help them achieve the most points in the allotted time. The game is visualized in Figure 2.

3.2.1 Game environment

The game takes place in a square-shaped room that contains several interactive elements distributed across its walls. One wall displays the high score list, showing the top nine scores. Another wall shows the current score, which resets at the beginning of each new round. A third wall presents the remaining time for the current round. The final wall contains a set of sliders that allow players to adjust the parameters of the interaction method in use.

3.2.2 Gameplay

The core mechanics of the game revolve around flying robots. The robots are randomly scattered throughout the room and could occlude each other. As soon as the robots have reached their position in the room for the round, the target or distractor appear in a sphere that makes up their body: The target is the letter "C" and the distractors are the letter "O". Players must identify and destroy the correct target using the specific gaze selection method for the current experimental block. Properly selecting the target from the distractors gives no penalty, whereas improperly selecting the distractor results in points subtracted from the score. When the correct target is

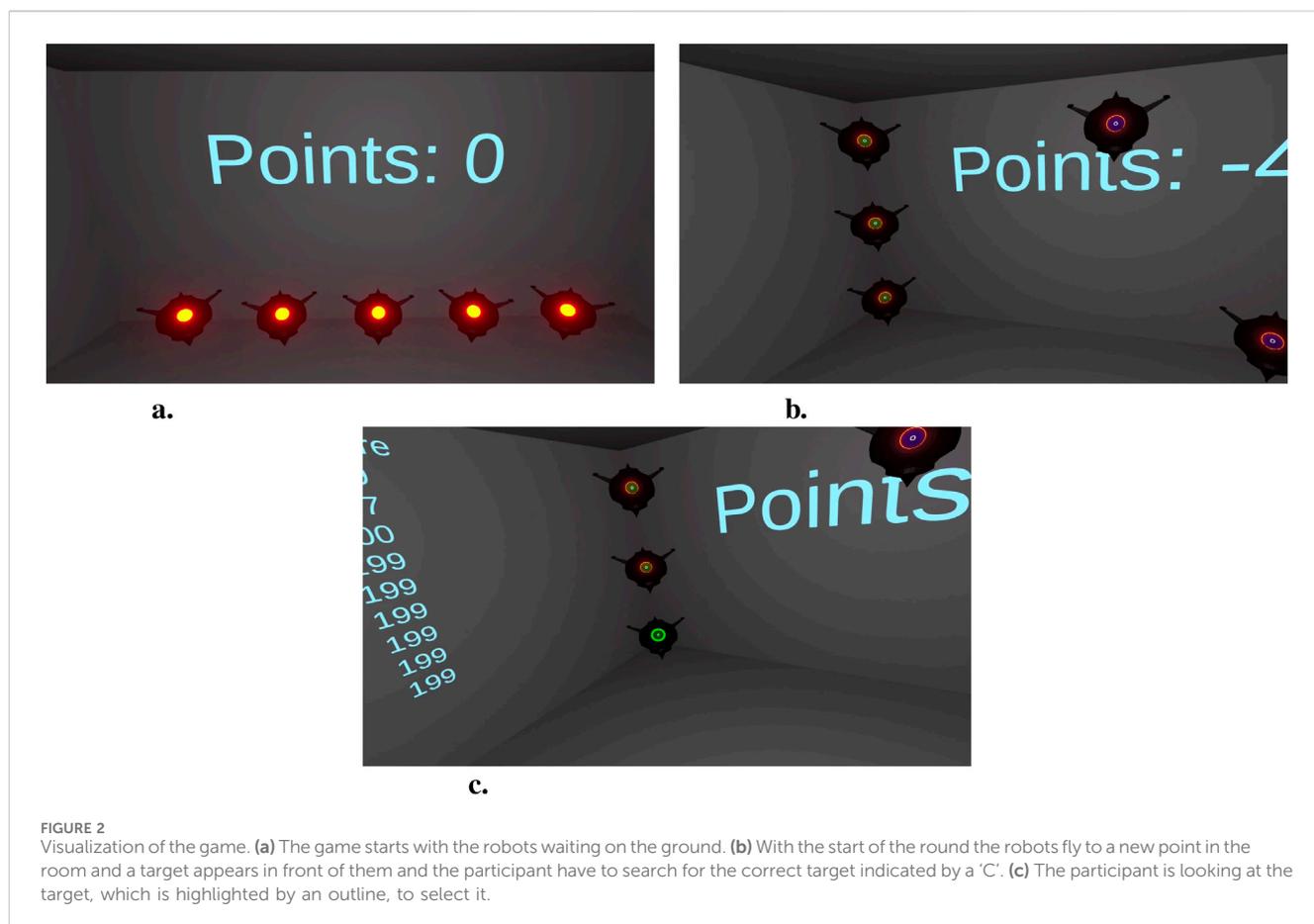


FIGURE 2

Visualization of the game. (a) The game starts with the robots waiting on the ground. (b) With the start of the round the robots fly to a new point in the room and a target appears in front of them and the participant have to search for the correct target indicated by a 'C'. (c) The participant is looking at the target, which is highlighted by an outline, to select it.

destroyed, the robots fly around to new positions and a new target has to be found. This continues for 30 s, then the round is over. The four experiment blocks are rounds related to each gaze selection method.

3.2.3 Scoring System

Points are awarded based on the player's performance according to the following rules: Positive points range from 5 to 20 for successfully destroying the correct target. If the player takes more than 10 s to destroy the target, they will only receive the minimum of five points. Otherwise, the points are determined through linear interpolation, with faster responses yielding higher scores. Minus points are awarded if an incorrect target is destroyed or a robot is shot at, resulting in a deduction of 21 points. This choice for scoring adds an additional time pressure to the participant.

3.3 VR setup

The game was conducted in VR. Participants interacted with the system using the HTC Vive Pro Eye (HTC Corporation, Taoyuan, Taiwan), which includes a built-in Tobii eye tracker (Core SW 2.16.4.67) with an estimated accuracy of $0.5^\circ - 1.1^\circ$ and a sampling frequency of 120 Hz. Eye tracking data was calibrated and accessed via the Vive SRanipal SDK (HTC Corporation, Taoyuan, Taiwan) (Sipatchin et al., 2021). In addition, the

participants were provided with an HTC Vive Pro Controller 2.0 to adjust the slider for the parameters for each interaction method and to answer the NASA-TLX questions in the questionnaire scene.

To extract gaze data from the VR headset, we employed the ZERO-Interface (Hosp and Wahl, 2023), which is integrated into VisionaryVR. This interface provides separate three-dimensional gaze vectors for each eye, along with a combined gaze vector. The data are accessible in real time, allowing them to be used for both gameplay and executing interaction methods. Additionally, all gaze data were recorded for further analysis.

3.4 Participants

52 participants from the University of Tübingen and the surrounding area took part in this study. 32 of them self-identified as women, 18 identified as men, one identified as non-binary, and one preferred not to provide gender information. Table 1 also shows the age range and how many had experience with VR and ET. We defined experience as having at least once used a VR or ET device. We did not specify level of understanding of the devices capabilities.

This study was reviewed and approved by Faculty of Medicine at the University of Tübingen with a corresponding ethical approval identification code 986/2020BO2. Participants provided their written informed consent to participate in this study.

TABLE 1 Demographic and experience data of study participants, including total number, average age, age range (minimum and maximum age) and experience with VR and ET technologies. The data is presented in total and broken down by gender identity to provide insights into the diversity of participants and corresponding technological familiarity.

	Count	Mean age	Minimum age	Maximum age	VR experience	ET experience
Women	32	26.31	18	63	15	11
Men	18	27.56	18	67	10	2
Non-binary	1	24.0	24	24	0	1
Prefer Not to Say	1	23.0	23	23	0	0
General	52	26.63	18	67	25	14

3.5 Study design

At the beginning of the experiment, participants were introduced to the game (similar to Section 3.2) and instructed on how to use the controller to make settings and start a round. For each condition, participants first played the game and adjusted the settings as required. Once they were satisfied with the configuration, they continued playing until they had completed ten consecutive rounds with no further changes. These ten rounds were later also used for the analysis. After completing the rounds, the NASA-TLX questionnaire was presented and participants provided their responses. This process was repeated for each condition. The first condition was always *Gaze Dwell*, as this is the simplest method that allows participants to familiarize themselves with the game mechanics. The order of the remaining conditions was randomized to minimize order effects.

3.6 Measurements

Several measurements were carried out to evaluate and compare the methods. In addition to the NASA-TLX questionnaire, participants were asked which method they would prefer if they had to choose one. Objective measures, such as score, were also recorded for quantitative comparisons. Participants could change settings as often as they wished, but each gaze selection method required at least ten consecutive rounds at their fixed settings. Objective measurements were only taken in these ten rounds to determine the most comfortable settings for each participant.

3.6.1 NASA-TLX

That questionnaire was created to measure the task load of a participant. It is widely used and has six different dimensions: *Mental Demand* evaluates how much mental and perceptual effort a task requires. The participants were asked, how much of thinking and deciding they had to do. *Physical Demand* assesses the physical effort needed to complete the task, which included movement. *Temporal Demand* considers the time pressure to which the participant was exposed and determines whether they perceived the pace of work as hurried or leisurely. *Performance* evaluates how the participants felt the successfulness of the completion of the task. *Effort* measures the amount of physical and mental energy that the participants believe they had to expend to solve the task. *Frustration* analyses emotional stress and takes into account factors such as anger, uncertainty, and frustration during

the task. For each dimension, participants rate on a scale from low to high, except for performance, where it ranges from perfect to failure.

3.6.2 Objective measurements

In addition to the subjective ratings, several objective characteristics were measured directly from the data: *Time on Task* refers to the time it took participants to select the correct target. *Points* represent the scores that participants achieved in each round. *Fails* indicates the number of incorrect selections, i.e., selecting a sphere with the character "O" instead of the "C". Therefore, we have the number of points and errors per round of a participant, and the time for the task depends on the number of correct selections.

3.7 Data analysis

As the data did not fulfill the assumptions required for ANOVA or t-tests, we used non-parametric tests. In order to compare all groups, the Kruskal–Wallis test (Kruskal and Allen, 1952; McKnight and Najab, 2010) was used, which, like ANOVA, tests whether there are significant differences between several groups, but does not assume normality or homogeneity of variances. It tests whether the median values of the groups differ significantly. The Mann-Whitney U test (Mann and Whitney, 1947; McKnight and Najab, 2010), a non-parametric alternative to the t-test, was used to compare two groups. It assesses whether two independent groups have significantly different distributions without assuming normality. The use of these tests ensures valid statistical conclusions given the non-normality of our data.

4 Results

We first evaluate the subjective measurements, namely, participant preferences and NASA-TLX responses regarding each interaction method. We then follow with the objective measurements like task duration and performance related to interaction method. As we are interested in which settings they preferred for each interaction method, we then compare the distributions of each setting.

4.1 Subjective measurements

Overall, the *Nod* method proved to be the most favored selection method by the participants, followed by the *head*

TABLE 2 The preferences of the participants for different methods, indicating the number of people who selected each method as their preferred method. Preferences are categorized by gender identity for the four methods.

	Gaze dwell	Head and gaze	Nod	Smooth pursuit
Women	10	6	13	3
Men	2	8	8	0
Non-Binary	1	0	0	0
Prefer Not to Say	0	1	0	0
General	13	15	21	3

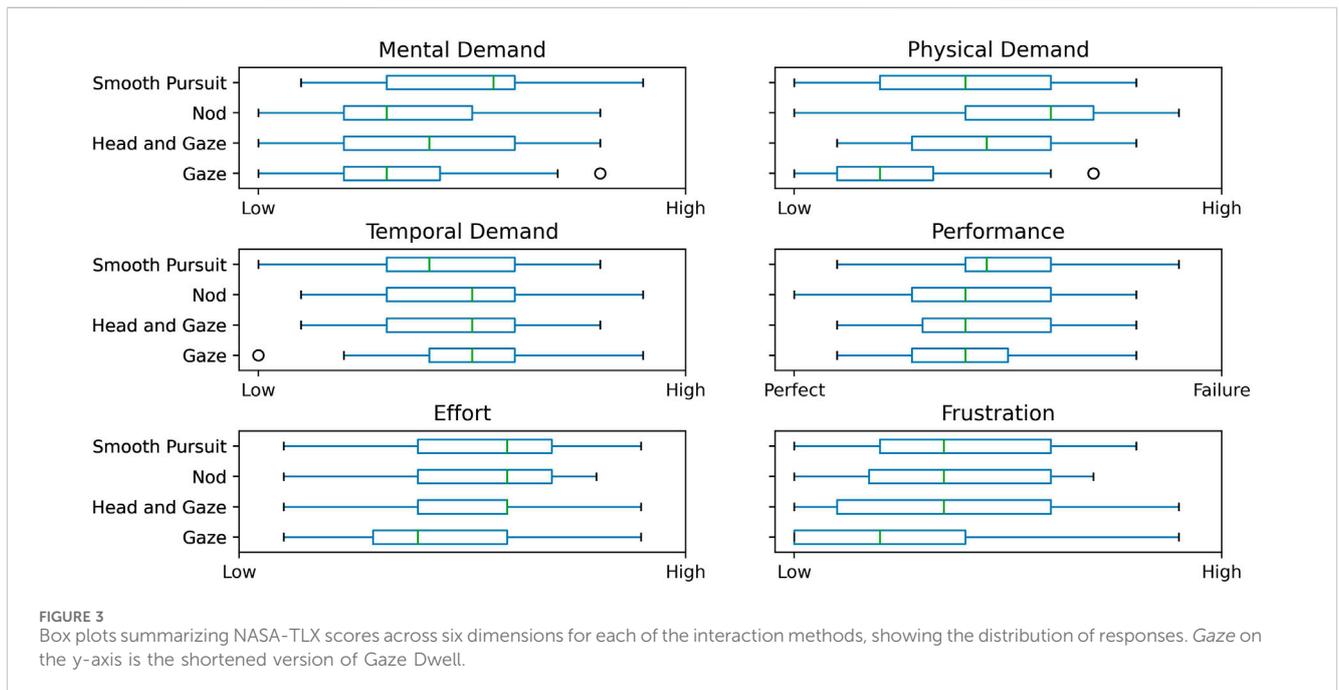


FIGURE 3 Box plots summarizing NASA-TLX scores across six dimensions for each of the interaction methods, showing the distribution of responses. Gaze on the y-axis is the shortened version of Gaze Dwell.

and gaze and Gaze Dwell methods, which showed a similar level of popularity. Table 2 details the preferences for each of the interaction methods. We observed unexpected differences in preferences across gender categories, especially for the Gaze Dwell method, which is predominantly favored by women participants: Ten out of a total of 13 who preferred this method. In contrast, men participants showed equal preference between the methods Head and Gaze and Nod, with eight participants preferring each of these options. The Smooth Pursuit method showed a minimal preference in all groups, with only three participants - all women - choosing this method. Overall, these results emphasize the broad appeal of the Nod method while showing gender preferences within other methods.

For some NASA-TLX dimensions, there were slight overall differences between the methods (see Figure 3). In particular, there are statistically significant differences in the dimensions of Mental Demand, Physical Demand, Effort and Frustration, with Smooth Pursuit and the other gaze-based interaction methods showing significant differences across multiple measures. We report these in detail below.

4.1.1 Mental demand

The methods show significant differences between the groups in these dimensions ($H = 10.694, p = 0.014$). The *post hoc* Mann-Whitney U test showed that the Smooth Pursuit method is significantly different from the other methods ($p < 0.05$) and with the information from Figure 3 we see that it is higher. No significant differences were found for the other methods ($p > 0.05$).

4.1.2 Physical demand

The physical demand dimension has a significant difference between the methods ($H = 29.720, p < 0.001$). The Gaze Dwell method is particularly low, while the others lead to a higher physical demand, with Nod reaching the highest value. Only between Smooth Pursuit and Head and Gaze no significant difference was found ($U = 1507.0, p = 0.311$).

4.1.3 Temporal demand

No statistically significant differences were found between the methods in this dimension ($H = 6.502, p = 0.09$). However, it is noteworthy that the median temporal demand in the Smooth

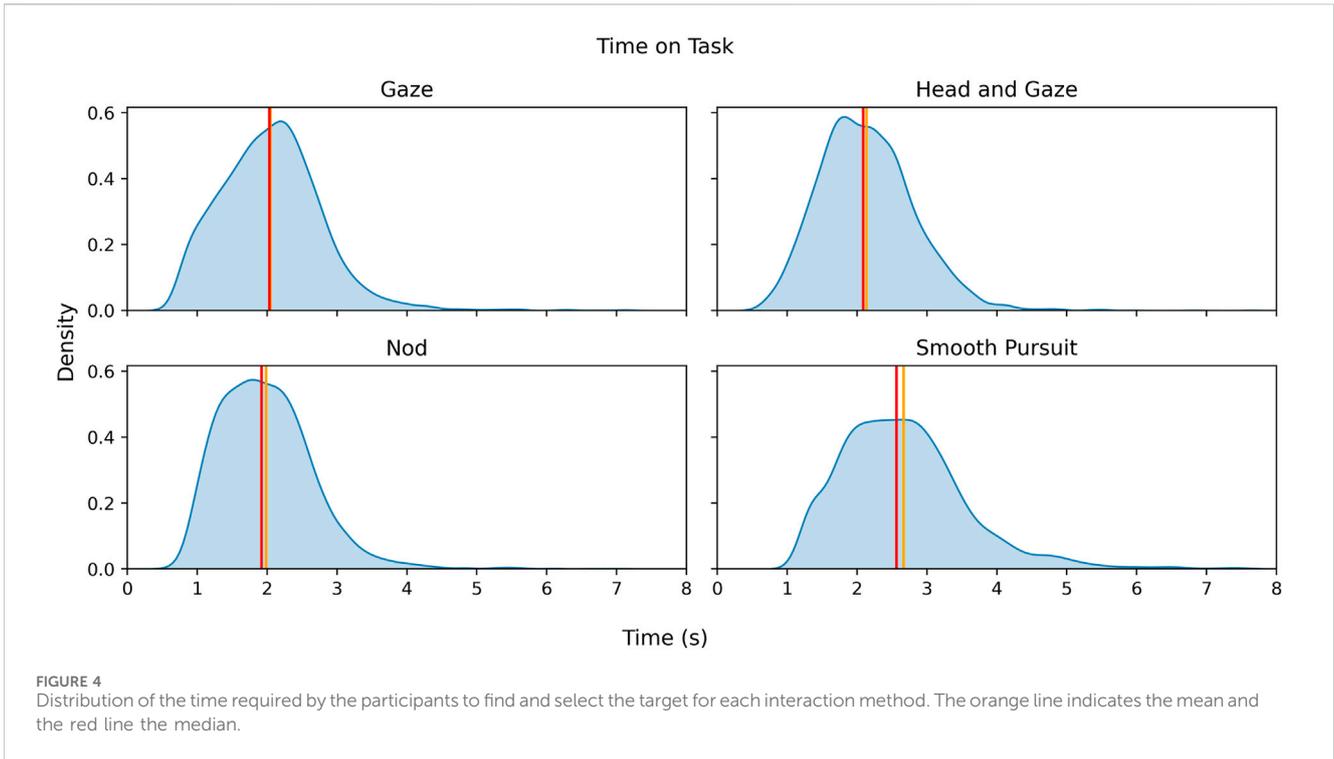


TABLE 3 Summarized performance metrics for the methods. The mean time is the average time it takes participants to find the correct target. The number of incorrect targets is the number of incorrectly selected targets over all rounds of the methods and the points are the average value over all rounds.

	Mean time (s)	Std time	Fails	Mean points	Std points
Gaze Dwell	2.049	0.716	171	144.626	32.131
Head and Gaze	2.137	0.678	73	147.737	25.573
Nod	1.985	0.691	114	159.578	29.265
Smooth Pursuit	2.665	0.905	21	122.810	23.150

Pursuit method is lower compared to the other methods, indicating a possible tendency for this approach to cause less time pressure.

4.1.4 Performance

No significant differences were found in this dimension either ($H = 7.257, p = 0.064$). It is noteworthy that the *Smooth Pursuit* method tends to perform below average compared to the other methods, which indicates a possible tendency towards lower effectiveness, especially compared to the gaze and head nodding methods.

4.1.5 Effort

There are significant differences between the methods based on the Kruskal test ($H = 11.308, p = 0.010$). The *post hoc* test showed that only the gaze differs significantly from the other methods ($p < 0.05$). The effort is therefore significantly lower than with the other methods. The other methods showed no significant differences ($p > 0.05$).

4.1.6 Frustration

This is similar to the effort dimension. There are significant differences between the methods ($H = 10.917, p = 0.012$), but these are between Gaze Dwell and the others. The perceived frustration is lower with the Gaze dwell method than with the other methods.

4.2 Objective measurements

Figure 4 shows the distribution of the time the participants take to destroy the target. There are significant differences between the interaction methods ($H = 1803.864, p < 0.001$). A *post hoc* Mann-Whitney-U test with adjusted p-values using the Bonferroni method shows that these differences are between all groups ($p < 0.001$).

Table 3 shows summarized performance metrics for each method, including the average score, the time to find the correct target, and the number of incorrect target choices across all rounds. Significant differences in scores between methods are observed ($K = 550.302, p < 0.001$). The significant differences between the

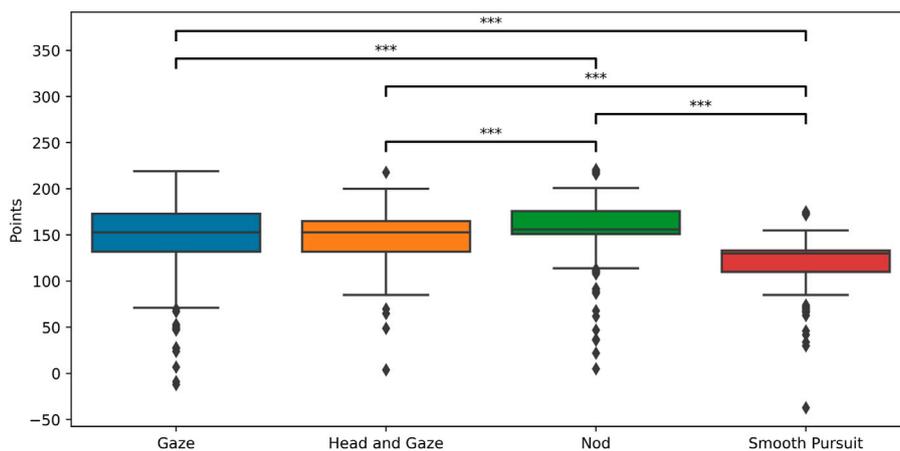


FIGURE 5
Box plot summary of the points obtained with the different interaction methods. Significant differences between the methods are labelled with ‘***’ to indicate $p < 0.001$.

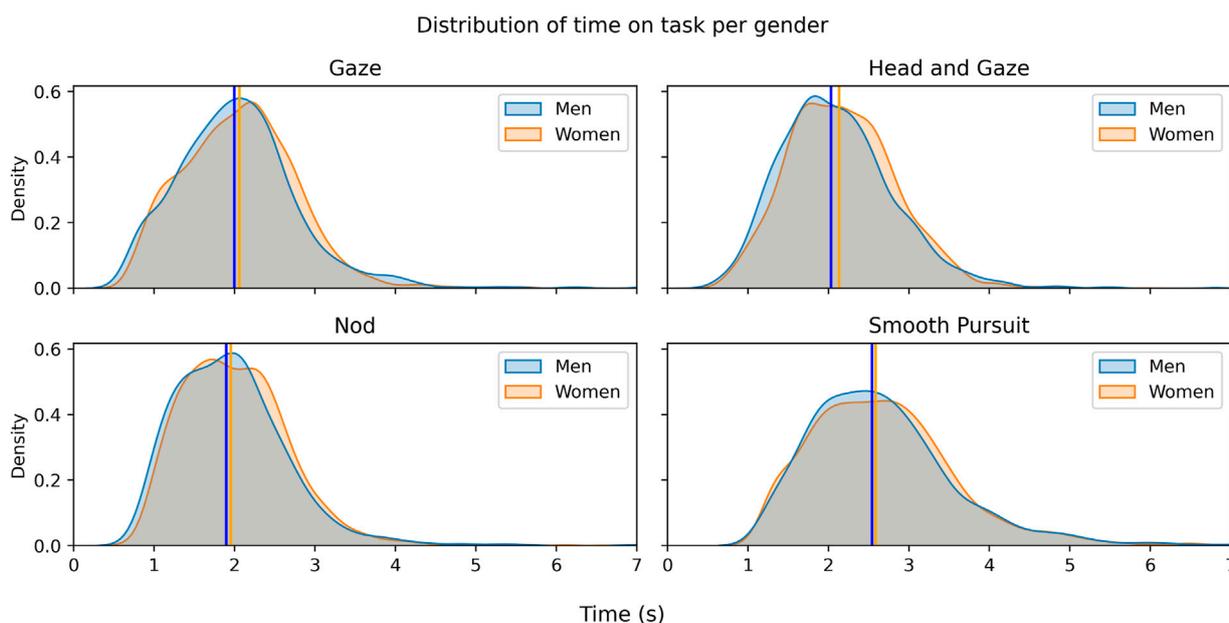


FIGURE 6
Distribution of time taken to select target, broken down by gender. The vertical lines indicate the median value.

methods are illustrated in **Figure 5**. *Gaze Dwell* and *Head and Gaze* score similar points, while *Nod* scores were the highest and *Smooth Pursuit* scores were the lowest.

4.3 Gender differences

Since preference for the interaction method suggests that gender has a potential influence, we decided to further examine whether gender differences were also apparent in NASA-TLX questionnaire results. Overall, we observed no significant differences between

genders across the NASA-TLX dimensions ($p > 0.05$). As, we only had two participants who identified as non-binary and preferred not to disclose this information, we do not wish to over-generalize their feedback.

Figure 6 illustrates the distribution of the time required to select the correct targets, separated by gender. The total time to find the correct target shows only small differences between the genders and the different methods. However, significant differences are observed for all methods, except for *Smooth Pursuit*, where times did not differ significantly between genders (Mann-Whitney U-test, $p < 0.05$ for other methods). This suggests that although selection

times are generally comparable, subtle but statistically significant gender differences can be observed for certain methods.

Table 4 presents performance metrics by gender. While no significant differences are found in points scored (all $p > 0.05$), significant differences are apparent in the time taken to select the correct target. Men generally make more incorrect selections than women, particularly with the *Nod* method ($U = 32927.0$, $p = 0.015$), and also show slightly higher error rates in *Smooth Pursuit* ($U = 32752.5$, $p = 0.004$). For *head and gaze*, men also show a higher error rate, but this does not differ significantly from women error rates ($U = 34486.0$, $p = 0.164$). In contrast, only with the *Gaze Dwell* method do women make more errors than men, but this was not significant ($U = 29251.5$, $p = 0.493$).

As our methodology allowed for participants to choose their thresholds for each method time, we can see which customization options could affect preference and performance. Figure 7 shows the distribution of the parameters used for each method, broken down by gender. Statistically significant differences between the genders are only found for the *Nod* parameters, especially for *Nod Strength* ($U = 143.0$, $p = 0.003$) and *Nod Direction Precision* ($U = 191.5$, $p = 0.048$). However, the distributions for time to select for both *head and gaze* and *gaze dwell* show that women may prefer slightly longer selection times than men. For the *Time to Select*, no significant differences were observed between the *Gaze Dwell* and *Head and Gaze* methods, for both men ($U = 182.5$, $p = 0.525$) and women participants ($U = 559.0$, $p = 0.528$). The lack of significant differences in these cases could be attributed to the different sample sizes of the groups. Also, there were no significant differences between groups for head alignment precision for the *head and gaze methods*, but, interestingly, we see from the distributions of both groups, that a few participants actually tolerated very low precision (over 17 degrees offset). Thus, apart from the *nod* parameters, the selection settings were generally consistent across genders. However, certain trends regarding some parameters allude to potential gender and personalization effects.

4.4 Influence of preferred method on objective metrics

Figure 8 shows that objective measures, such as the number of points achieved and the time to select, do not differ significantly between participants who favored a method and those who did not. However, there is a significant difference in the scores for the *Nod* method ($U = 43554.0$, $p < 0.001$), while the time to select for the *Head and Gaze* ($U = 26778.5$, $p = 0.0012$) and *Nod* ($U = 24843.0$, $p < 0.001$) methods is significantly different.

Further analysis of the NASA-TLX responses revealed no significant differences between participants who favored a method and those who did not. This suggests that subjective perceptions of workload, such as mental and physical demands, effort and frustration, are relatively consistent regardless of the preferred method.

5 Discussion

For this study, we developed a VR game environment to evaluate and compare four common gaze-based interaction methods,

TABLE 4 Performance metrics by gender for all four methods, showing mean points scored, mean time to select the correct target and total number of failures per participant across all rounds.

Method	Gender	Points	Time (s)	Error rate
Gaze Dwell	Men	144.23	2.03	3.11
	Women	144.17	2.06	3.53
Head and Gaze	Men	148.47	2.10	1.61
	Women	146.20	2.17	1.34
Nod	Men	159.38	1.98	3.11
	Women	158.72	2.03	1.69
Smooth Pursuit	Men	121.70	2.68	0.78
	Women	122.86	2.67	0.19

focusing on subjective and objective measures of user experience and performance. Using NASA-TLX questionnaire, subjective workload was assessed across six dimensions, while objective measures such as time on task, score, and error rate provided quantitative insights. Participants adjusted interaction settings for optimal comfort, and results were analyzed for possible gender differences in preferences, parameter adjustment, and performance. This approach aimed to identify the strengths and weaknesses of each method and provided valuable insights for designing more intuitive and user-friendly gaze-based interfaces.

The NASA-TLX results show significant differences in several dimensions between the interaction methods. The smooth pursuit for target selection in particular showed a higher mental demand than other methods, indicating a potentially more cognitively intensive interaction. The physical demand varied significantly, with the gaze method requiring the least physical effort. Although the overall differences in time requirements were not statistically significant. Performance ratings showed only marginal differences, with smooth pursuit tending to lead to lower perceived success, especially compared to gaze dwell and nodding movements. Effort and frustration scores were significantly lower for the gaze dwell method compared to the other methods, which can be attributed to the ease of use and lower frustration scores. More importantly, these results were consistent between genders and there were no significant gender differences in perceived workload. These results emphasize the different cognitive and physical profiles of each method, with the gaze dwell method proving to be the most user-friendly in terms of physical and emotional demands, while the smooth pursuit method had higher demands on mental and time resources.

The performance metrics show significant differences between the interaction methods in terms of time, accuracy, and score. The nod method achieved the highest average score, while the smooth pursuit method had the lowest average score but resulted in the fewest incorrect selections, indicating a trade-off between speed and accuracy. The gaze dwell time and head and gaze methods achieved comparable results, although the gaze dwell time method produced the highest number of incorrect selections. The average time to complete the task varied considerably, with the smooth pursuit method taking the longest time overall. These results highlight clear performance conflicts where the nod method

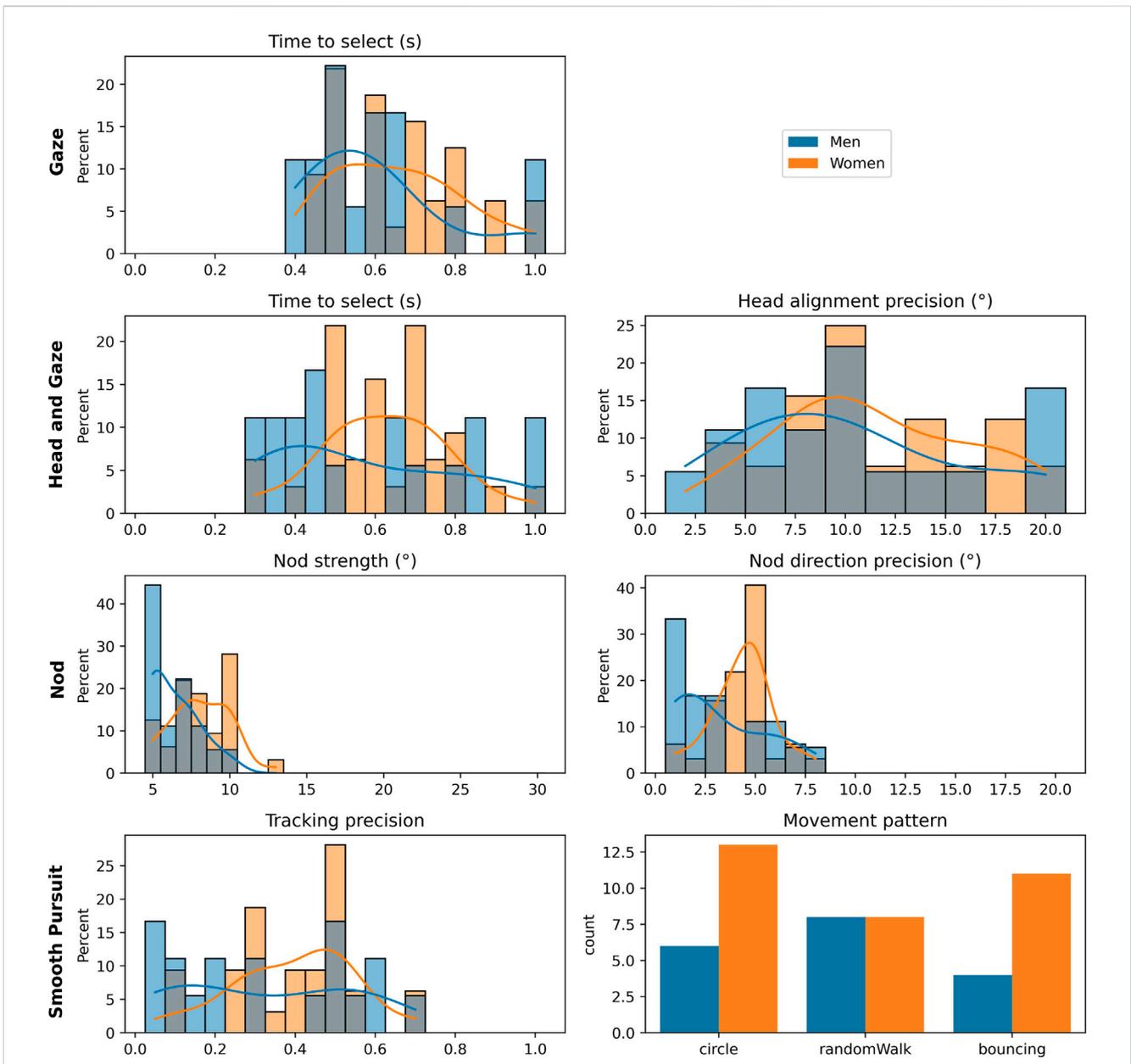


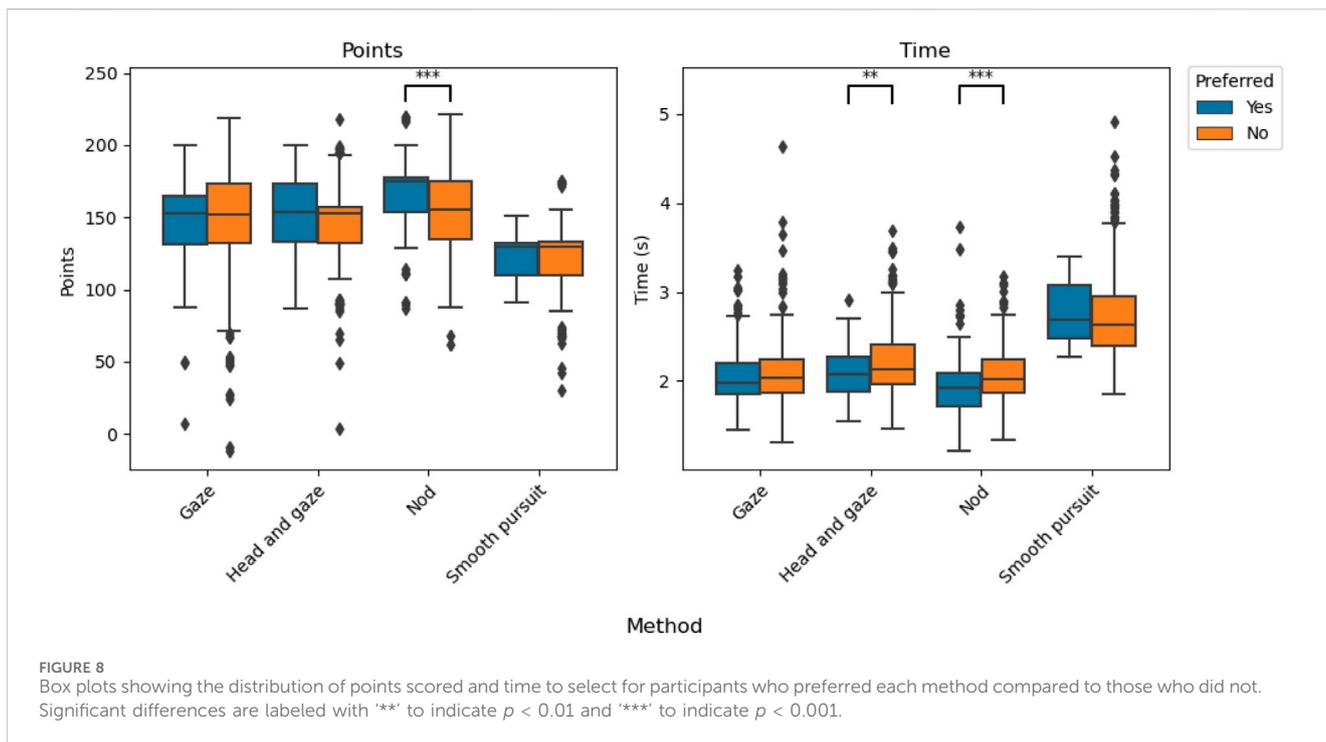
FIGURE 7 Histogram of selected parameters per gender. The first line is for the *Gaze Dwell* method, the second line for the *Head and Gaze* method, the third for the *Nod* method and the last for *Smooth Pursuit*.

maximizes scoring potential, but the smooth pursuit is more error resistant.

Regarding preference, the nod interaction method was the most preferred in general, with men preferring mainly the head and gaze or nod methods and women preferring the gaze dwell or nod methods. However, the analyses of the NASA-TLX questionnaire show no significant differences in the perceived task load between the genders. There are significant differences in the objective measurements, but these are so small that they are not recognized in reality. This suggests that the observed preferences for the gaze dwell method are not due to differences in task demands or workload. Instead, they could stem from other factors, such as individual familiarity with gaze-based interactions, comfort level or

specific engagement with the game mechanics, suggesting a more nuanced understanding of the influence of gender on interaction preferences.

Gender differences in selection time and scoring are minimal, with both groups showing similar performance on all methods. However, we found that males tended to make more selection errors than females, particularly when using the *Nod* and *Smooth Pursuit* method. Significant gender differences in parameter preference selection occur only for the Nod Strength and Nod Direction Precision settings, while selection settings such as Time to Select remain consistent across genders for all gaze-based methods. These results suggest that while overall performance is comparable, subtle differences in error rates and parameter preferences indicate



different interaction needs between genders. This highlights the importance of designing more adaptive and inclusive VR interfaces that address to different users. Taking these gender tendencies into account when developing gaze-based selection methods could lead to more accessible and user-friendly experiences that are more responsive to individual needs.

Interestingly enough, though target selection through dwell time has been one of the most common gaze-based interaction methods (Narnakani et al., 2023; Chen and Shi, 2017) to combat the *Midas Touch* problem, we found other methods were only slightly more preferred. We attribute this to the structure of the game, which requires frequent and rapid target selection, the participants’ gaze is constantly in an intentional target selection mode, with little chance for unintentional eye movements, reducing the effect of the *Midas Touch* problem (Hyrskykari et al., 2012). Although the problem is less apparent compared to a more naturalistic task, there were still error rates, especially for the gaze dwell method. This suggests that while gaze dwell provides an intuitive and fast selection process, the associated unintended activations can lead to increased inaccuracies, especially in fast or continuous selection tasks. Potentially more naturalistic tasks may have users preferring other methods even more.

6 Potential implications

Our research aim was initially focused on how user preference for different gaze-based interaction paradigms related to personal opinions of demand and comfort, and how this could affect task performance factors. We did find that subjective comfort can sometimes come with a tradeoff of accuracy. Even more so, we did not expect gender to become such a relevant factor in our analyses. Therefore, we feel this research does have potential

implications for equitable design choices. We suggest that designing user interfaces that employ gaze should not only understand the importance of customizability and personal preference, but a user’s choice could be affected by other factors such as gender, diversity, or other sociodemographics. By no means are we implying digital tools for specific genders should be different—for instance, less precise and more pink for women—rather encouraging a broader perspective when developing these tools for varied use cases.

7 Limitations

While our study provides valuable insights into gaze-based interaction methods, it is important to note that these results come from a gaming environment. In such an environment, convenience and ease of interaction are often prioritized over strict efficiency, as users are generally more tolerant of occasional missteps if it improves their overall experience. This prioritization can differ significantly in real life or professional applications, where accuracy and efficiency are often paramount. Therefore, the trade-off between effort and efficiency observed in our study may not directly translate to contexts other than games. In practical applications where the stakes are high, users may prefer precise control and minimal errors over convenience, shifting the balance between these factors. Another important limitation of our study is the persistent problem of unintended selection, commonly referred to as the *Midas Touch* problem. Although we tested several confirmation techniques to mitigate this problem, false activations were still observed. This suggests that current gaze-based selection techniques are not yet optimally accurate, especially in fast or complex interaction scenarios. Future research should explore adaptive filtering techniques to reduce unintended selections. In addition, studies should test these methods in

scenarios with practical, real-world requirements to better assess their broader applicability.

8 Conclusion

This study investigated different gaze-based and combined gaze-head interaction methods in an interactive game environment and examined how these methods affect performance and error rates. Our results show significant differences between the methods in terms of perceived workload and accuracy, suggesting that some approaches provide intuitive and comfortable interaction, while others have higher precision. Gender differences were also found, with preferences and performance differing for certain methods, suggesting that interaction systems should take individual differences into account by offering customizable, personalisable options.

While gaze-only methods remain popular, combined gaze-head approaches show the potential for more accurate target selection with fewer errors. However, in a gaming environment, precision may not be as important to the participant, as errors have less severe consequences than in the real world, which may mean that these results are different in real-world applications. This emphasizes the importance of tailoring interaction methods to specific use cases, taking into account both preferences and performance requirements. Future research should further investigate these methods in different environments to test their practicality and inclusivity in real-world scenarios.

Data availability statement

The datasets presented in this article are not readily accessible to ensure responsible use and compliance with ethical guidelines. Requests for access to the datasets should be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the Faculty of Medicine at the University of Tübingen with a corresponding ethical approval identification code 986/2020BO2. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

BS: Conceptualization, Data curation, Formal Analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review and editing. YS: Conceptualization,

Methodology, Software, Validation, Visualization, Writing – review and editing. AN: Conceptualization, Software, Writing – review and editing. RA: Conceptualization, Writing – review and editing. NC: Conceptualization, Methodology, Supervision, Validation, Visualization, Writing – review and editing. SW: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research is supported by European Union's Horizon 2020 research and innovation program under grant agreement No. 951910 and the German Research Foundation (DFG): SFB 1233, Robust Vision: Inference Principles and Neural Mechanisms, TP TRA, project No. 276693517.

Acknowledgments

We acknowledge support from the Open Access Publication Fund of the University of Tübingen.

Conflict of interest

Authors YS, NC, and SW were employed by Carl Zeiss Vision International GmbH.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Best, D. S., and Duchowski, A. T. (2016). "A rotary dial for gaze-based pin entry," in *Proceedings of the ninth biennial ACM symposium on eye tracking research and applications* (New York, NY, USA: Association for Computing Machinery), 69–76. doi:10.1145/2857491.2857527

Blattgerste, J., Renner, P., and Pfeiffer, T. (2018). "Advantages of eye-gaze over head-gaze-based selection in virtual and augmented reality under varying field of views," in *Proceedings of the workshop on communication by gaze interaction* (New York, NY, USA: Association for Computing Machinery). doi:10.1145/3206343.3206349

- Bovo, R., Binetti, N., Brumby, D. P., and Julier, S. (2020). "Detecting errors in pick and place procedures: detecting errors in multi-stage and sequence-constrained manual retrieve-assembly procedures." in *Proceedings of the 25th international conference on intelligent user interfaces*, 536–545.
- Bremer, G., Stein, N., and Lappe, M. (2021). "Predicting future position from natural walking and eye movements with machine learning," in *2021 IEEE international conference on artificial intelligence and virtual reality (AIVR)*, 19–28. doi:10.1109/AIVR52153.2021.00013
- Chen, Z., and Shi, B. E. (2017). Improving gaze-based selection using variable dwell time. *ArXiv abs/1704.06399*. doi:10.48550/arXiv.1704.06399
- Dey, P. P., Sinha, B. R., Amin, M., and Badkoobehi, H. (2019). Best practices for improving user interface design. *Int. J. Softw. Eng. and Appl.* 10, 71–83. doi:10.5121/ijsea.2019.10505
- Duchowski, A. T. (2018). Gaze-based interaction: a 30 year retrospective. *Comput. and Graph.* 73, 59–69. doi:10.1016/j.cag.2018.04.002
- Esteves, A., Velloso, E., Bulling, A., and Gellersen, H. (2015). "Orbits: gaze interaction for smart watches using smooth pursuit eye movements," in *Orbits: gaze interaction for smart watches using smooth pursuit eye movements*, 15. New York, NY, USA: Association for Computing Machinery, 457–466. doi:10.1145/2807442.2807499
- Fernandes, A. S., Schütz, I., Murdison, T. S., and and, M. J. P. (2025). Gaze inputs for targeting: the eyes have it, not with a cursor. *Int. J. Human-Computer Interact.* 0, 1–19. doi:10.1080/10447318.2025.2453966
- Gou, Z., Shao, Z., Gong, Y., Shen, Y., Yang, Y., Duan, N., et al. (2023). Critic: large language models can self-correct with tool-interactive critiquing. *arXiv preprint arXiv:2305.11738*
- Gür, D., Schäfer, N., Kupnik, M., and Beckerle, P. (2020). A human-computer interface replacing mouse and keyboard for individuals with limited upper limb mobility. *Multimodal Technol. Interact.* 4, 84. doi:10.3390/mti4040084
- Haas, J. K. (2014). A history of the unity game engine
- Hansen, J. P., Tørning, K., Johansen, A. S., Itoh, K., and Aoki, H. (2004). "Gaze typing compared with input by head and hand," in *Proceedings of the 2004 symposium on eye tracking research and applications* (New York, NY, USA: Association for Computing Machinery), 131–138. doi:10.1145/968363.968389
- Hart, S. G., and Staveland, L. E. (1988). "Development of nasa-tlx (task load index): results of empirical and theoretical research," in *North-holland, vol. 52 of Advances in psychology*. Editors H. Mental Workload, P. A. Hancock, and N. Meshkati, 139–183. doi:10.1016/S0166-4115(08)62386-9
- Hoanca, B., and Mock, K. (2006). "Secure graphical password system for high traffic public areas," in *Proceedings of the 2006 symposium on eye tracking research and applications* (New York, NY, USA: Association for Computing Machinery), 35, 35. doi:10.1145/1117309.1117319
- Hosp, B. W., Dechant, M., Sauer, Y., Severitt, B., Agarwala, R., and Wahl, S. (2024). Visionaryvr: an optical simulation tool for evaluating and optimizing vision correction solutions in virtual reality. *Sensors* 24, 2458. doi:10.3390/s24082458
- Hosp, B. W., and Wahl, S. (2023). "Zero: a generic open-source extended reality eye-tracking controller interface for scientists," in *Proceedings of the 2023 symposium on eye tracking research and applications* (New York, NY, USA: Association for Computing Machinery). doi:10.1145/3588015.3589203
- Huckauf, A., and Urbina, M. (2007). "Gazing with peye: new concepts in eye typing," in *Proceedings of the 4th symposium on applied perception in graphics and visualization*, 141. New York, NY, USA: Association for Computing Machinery. doi:10.1145/1272582.1272618
- Huckauf, A., and Urbina, M. H. (2008). "Gazing with peyes: towards a universal input for various applications," in *Proceedings of the 2008 symposium on eye tracking research and applications* (New York, NY, USA: Association for Computing Machinery), 51–54. doi:10.1145/1344471.1344483
- Hyrskykari, A., Istance, H., and Vickers, S. (2012). "Gaze gestures or dwell-based interaction?," in *Proceedings of the symposium on eye tracking research and applications* (New York, NY, USA: Association for Computing Machinery), 12, 229–232. doi:10.1145/2168556.2168602
- 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* (New York, NY, USA: Association for Computing Machinery), 90, 11–18. doi:10.1145/97243.97246
- Jha, S., Seo, C., Yang, E., and Joshi, G. P. (2021). Real time object detection and tracking system for video surveillance system. *Multimedia Tools Appl.* 80, 3981–3996. doi:10.1007/s11042-020-09749-x
- Kruskal, W. H., and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* 47 (260), 583–621. doi:10.1080/01621459.1952.10483441
- Lee, K. F., Chen, Y. L., Yu, C. W., and Wu, C. H. (2020). "The eye tracking and gaze estimation system by low cost wearable devices," in *2020 IEEE international conference on consumer electronics - taiwan (ICCE-Taiwan)*, 1–2. doi:10.1109/ICCE-Taiwan49838.2020.9258009
- Lin, Z., Trivedi, S., and Sun, J. (2023). Generating with confidence: uncertainty quantification for black-box large language models. *arXiv Prepr. arXiv:2305*.
- Macaranas, A., Antle, A. N., and Riecke, B. E. (2015). What is intuitive interaction? Balancing users' performance and satisfaction with natural user interfaces. *Interact. Comput.* 27, 357–370. doi:10.1093/iwc/iwv003
- Majoranta, P., and Bates, R. (2009). Special issue: communication by gaze interaction. *Univers. Access Inf. Soc.* 8, 239–240. doi:10.1007/s10209-009-0150-7
- Majoranta, P., and Riih a, K.-J. (2002). "Twenty years of eye typing: systems and design issues," in *Proceedings of the 2002 symposium on eye tracking research and applications* (New York, NY, USA: Association for Computing Machinery), 15–22. doi:10.1145/507072.507076
- Mann, H. B., and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statistics* 18, 50–60. doi:10.1214/aoms/1177730491
- Mastrangelo, A. S., Karkhanis, M., Likhite, R., Bulbul, A., Kim, H., Mastrangelo, C. H., et al. (2018). "A low-profile digital eye-tracking oculometer for smart eyeglasses," in *2018 11th international conference on human system interaction (HSI) IEEE*, 506–512.
- McKnight, P. E., and Najab, J. (2010). *Kruskal-wallis test*. John Wiley and Sons, Ltd, 1. doi:10.1002/9780470479216.corpsy0491
- McKnight, P. E., and Najab, J. (2010). *Mann-whitney U test*. John Wiley and Sons, Ltd, 1. doi:10.1002/9780470479216.corpsy0524
- 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* (New York, NY, USA: Association for Computing Machinery), 23, 1–17. doi:10.1145/3544548.3580871
- Pannasch, S., Helmert, J. R., Malischke, S., Storch, A., and Velichkovsky, B. M. (2008). Eye typing in application: a comparison of two systems with als patients. *J. Eye Mov. Res.* 2. doi:10.16910/jemr.2.4.6
- Pelz, J. B., and Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vis. Res.* 41, 3587–3596. doi:10.1016/S0042-6989(01)00245-0
- Pfeuffer, K., Mayer, B., Mardanbegi, D., and Gellersen, H. (2017a). "Gaze + pinch interaction in virtual reality," in *Proceedings of the 5th symposium on spatial user interaction* (New York, NY, USA: Association for Computing Machinery), 17, 99–108. doi:10.1145/3131277.3132180
- Pfeuffer, K., Mayer, B., Mardanbegi, D., and Gellersen, H. (2017b). "Gaze+ pinch interaction in virtual reality," in *Proceedings of the 5th symposium on spatial user interaction*, 99–108.
- Piumsomboon, T., Lee, G., Lindeman, R. W., and Billingham, M. (2017). Exploring natural eye-gaze-based interaction for immersive virtual reality. *2017 IEEE Symposium 3D User Interfaces 3DUI*, 36–39. doi:10.1109/3DUI.2017.7893315
- Plopski, A., Hirzle, T., Norouzi, N., Qian, L., Bruder, G., and Langlotz, T. (2022). The eye in extended reality: a survey on gaze interaction and eye tracking in head-worn extended reality. *ACM Comput. Surv.* 55, 1–39. doi:10.1145/3491207
- 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* (New York, NY, USA: Association for Computing Machinery), 91–98. doi:10.1145/3131277.3132182
- Rakhmatulin, I. (2020). A review of the low-cost eye-tracking systems for 2010–2020. *Corr. abs/2010.05480*.
- Sidenmark, L., and Gellersen, H. (2019). "Eye&head: synergetic eye and head movement for gaze pointing and selection," in *Proceedings of the 32nd annual ACM symposium on user interface software and technology* (New York, NY, USA: Association for Computing Machinery), 1161–1174. doi:10.1145/3332165.3347921
- Sipatchin, A., Wahl, S., and Rifai, K. (2021). Eye-tracking for clinical ophthalmology with virtual reality (vr): a case study of the htc vive pro eye's usability. *Healthcare* 9, 180. doi:10.3390/healthcare9020180
- Smith, J. D., and Graham, T. N. (2006). "Use of eye movements for video game control," in *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology*. 20–es.
- Špakov, O., and Majoranta, P. (2012). "Enhanced gaze interaction using simple head gestures," in *Proceedings of the 2012 ACM conference on ubiquitous computing* (New York, NY, USA: Association for Computing Machinery), 12, 705–710. doi:10.1145/2370216.2370369
- Starker, I., and Bolt, R. A. (1990). "A gaze-responsive self-disclosing display," 90. New York, NY, USA: Association for Computing Machinery, 3–10. doi:10.1145/97243.97245Proceedings of the SIGCHI conference on human factors in computing systems

Subramanian, M., Songur, N., Adjei, D., Orlov, P., and Faisal, A. A. (2019). "A. eye drive: gaze-based semi-autonomous wheelchair interface," in *2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)* IEEE, 5967–5970.

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* (New York, NY, USA: Association for Computing Machinery), 13, 439–448. doi:10.1145/2493432.2493477

Ward, D. J., and MacKay, D. J. (2002). Fast hands-free writing by gaze direction. *Nature* 418, 838. doi:10.1038/418838a

Wei, Y., Shi, R., Yu, D., Wang, Y., Li, Y., Yu, L., et al. (2023). "Predicting gaze-based target selection in augmented reality headsets based on eye and head endpoint distributions," in *Proceedings of the 2023 CHI conference on human factors in computing systems* (New York, NY, USA: Association for Computing Machinery). doi:10.1145/3544548.3581042

Wobbrock, J. O., Rubinstein, J., Sawyer, M. W., and Duchowski, A. T. (2008). "Longitudinal evaluation of discrete consecutive gaze gestures for text entry," in *Proceedings of the 2008 symposium on eye tracking research and applications* (New York, NY, USA: Association for Computing Machinery), 11–18. doi:10.1145/1344471.1344475