

Gaze-Based Menu Navigation in Virtual Reality: A Comparative Study of Layouts and Interaction Techniques

László Kopácsi[1(✉)], Albert Klimenko[2], Abdulrahman Mohamed Selim[1], Michael Barz[1,3], and Daniel Sonntag[1,3]

[1] Interactive Machine Learning, German Research Center for Artificial Intelligence (DFKI), Saarbrücken, Germany
{laszlo.kopacsi,abdulrahman.mohamed,michael.barz,daniel.sonntag}@dfki.de
[2] Saarland University, Saarbrücken, Germany
s8alklim@stud.uni-saarland.de
[3] Applied Artificial Intelligence, University of Oldenburg, Oldenburg, Germany

Abstract. Integrating eye-tracking technologies in Extended Reality (XR) headsets has enabled intuitive, hands-free system interaction, such as gaze-based menu navigation. However, there is a lack of comprehensive comparisons and consensus in the literature on the optimal use of gaze-based menu navigation. This paper presents a comparative analysis of gaze-based menu navigation in virtual environments, focusing on two common menu layouts: pie and list menus, with three interaction methods: gaze-based dwell, controller-based, and a multimodal approach combining gaze and controller inputs. We conducted a 19-participant within-subject study, measuring task completion time, error rate, usability, and user preference for each condition. The results indicate that while the pie layout was statistically faster and less erroneous than the list layout, novice users tend to favour list layouts. Furthermore, we found that users preferred the multimodal interaction method, despite its lower task completion times and higher error rates compared to controller-based navigation. Based on our findings, we offer design guidelines and recommendations for implementing gaze-based menu systems.

Keywords: Extended Reality (XR) · Gaze-based Interaction · Menu Navigation · Eye Tracking

1 Introduction

Eye tracking has long been regarded as an important input modality for natural, hands-free interaction [24,54]. This notion holds true to this day, especially with the integration of eye tracking into Extended Reality (XR) devices (e.g., HTC Vive XR Elite[1] and Apple Vision Pro[2]), which use it to facilitate user interaction

[1] https://www.vive.com/eu/product/vive-xr-elite/ (Accessed 17 Feb 2025).
[2] https://www.apple.com/apple-vision-pro/ (Accessed 17 Feb 2025).

© The Author(s) 2026
C. Ardito et al. (Eds.): INTERACT 2025, LNCS 16108, pp. 520–543, 2026.
https://doi.org/10.1007/978-3-032-04999-5_31

with digital content [13,49]. This usage of eye tracking falls under active gaze-based interaction, where a user intentionally uses their gaze to interact with and control a system [14]. Menu manipulation is one application in which users can interact with immersive environments by using their gaze to navigate and select options. However, despite the availability of open standards that facilitate development across different headsets, in the literature, there is no consensus on the effectiveness or optimal use of gaze-based interaction for menu manipulation.

In general, menus in XR can be categorised based on three main aspects: (1) Interaction Method, (2) Menu Layout, and (3) Menu Placement [20]. For gaze-based interaction methods, the constant activity of the eyes presents a significant challenge in distinguishing between exploratory gaze behaviour and intentional command-activating gaze behaviour, which is known as the Midas Touch problem [24]. Dwell-based gaze input (i.e., intentionally maintaining gaze fixation on a specific target for a predetermined duration known as dwell time to activate or select elements [46]) is often regarded as one of the simplest and most common gaze events [49]. While it is commonly used for selection, e.g., [23], it has also been investigated for other actions, such as scrolling [22]. To overcome the Midas Touch problem, long dwell times (e.g., 500 ms [43] or 300 ms [45]) are often used, but this raises the challenge of finding the best compromise between error rates and interaction times [44]. As a result, researchers have explored more advanced gaze-based events such as smooth pursuits [28] (i.e., the eye movement while following a moving object [19]) and border-crossing [21] (i.e., intentionally moving the gaze across a pre-defined boundary to enter an activation area, thereby triggering an action). Multimodal systems have also been explored, e.g., using gaze for pointing and a controller for selecting [44].

Regarding menu layouts, pie or radial menus (i.e., where items are arranged around the circumference of a circle at an equal radial distance from the centre [7]) are one of the most researched menu layouts for gaze-based interaction [2,30,53]. However, comparative studies by Monteiro et al. [42] and Lediaeva and LaViola [34] indicate that users tend to prefer traditional linear or list layouts over pie layouts. Additionally, Lediaeva and LaViola [34] found no significant difference in task completion time between pie and list layouts during single-level menu selection using gaze-based interaction. Lastly, menu placement can be broadly categorised into head-referenced, e.g. [36], body-referenced, e.g. [34], and world-referenced, e.g. [42], configurations. World-referenced placement has been identified as the most suitable option to ensure a proper and fair comparison of different menu layouts and interaction methods [34,42]. Therefore, we decided to focus on menu layouts and interaction methods, while keeping the menu placement constant, i.e., world-referenced.

To provide a comprehensive overview of how different menu layouts and interaction methods affect the overall user experience in XR, we conducted a within-subject study with 19 participants using six conditions in Augmented Reality (AR). The study evaluated two hierarchical menu layouts: *Pie Menu* and *List Menu*, across three interaction methods: a *Gaze-only* dwell-based method (i.e., a baseline for gaze-based interaction), a *Controller-only* method (i.e., the default

interaction method in XR), and a multimodal *Gaze-and-Controller* method in which gaze is used for pointing and the controller for selection. Although *Border-crossing* was initially included as an additional Gaze-only interaction method, it was removed prior to the main user study due to our findings in the pilot study. For each of the six conditions, we measured task completion time, error rate, usability, and user preference. Our contributions are twofold: (i) we conducted this extensive study to provide a comparison between the different menu and interaction combinations, and (ii) we formulated design guidelines based on the quantitative and post-experiment questionnaire results. Our results show that users had a slight preference towards the Gaze-and-Controller multimodal interaction method despite its lower task completion times and higher error rates compared to the Controller-only method. In addition, our findings indicate that Pie layouts are statistically faster and less erroneous than List layouts using Controller-only and Gaze-and-Controller interaction methods, but users without prior XR and eye tracking experience still preferred List menus. To support reproducibility and future research, our source code is available at https://github.com/DFKI-Interactive-Machine-Learning/menu-navigation-in-VR.

2 Background and Related Work

XR is an umbrella term for technologies that alter or generate reality. It encompasses AR, which overlays digital elements onto the real world; Mixed Reality (MR), which enables interaction between real and digital elements; and Virtual Reality (VR), which creates a fully immersive digital environment [51]. In this section, we review the literature on digital menu layouts and interaction methods in XR, with a particular focus on gaze-based related approaches.

2.1 Menu Layout

Various menu layouts have been proposed over the years. However, unique layouts are often associated with specific interaction methods (e.g., [15,52]), making them difficult to compare against other menu layouts and interaction methods. Therefore, we decided to focus on two classical layouts, i.e., list and pie, because they are not associated with specific interaction methods. List layouts usually display items in a straight line, either horizontally or vertically, and remain the most commonly used menu layout in VR [20]. Pie layouts, on the other hand, arrange items in a circular pattern around a central point, leveraging the natural range of motion of the human arm and eyes, which enables quick access and selection. Huckauf and Urbina [21] introduced the concept of pie layouts for gaze-based interaction and demonstrated their efficiency in a typing task. Since then, pie layouts have shown a dominant presence in gaze-based interaction literature [2,30,44,53]. However, Monteiro et al. [42] and Lediaeva and LaViola [34] suggest that users still prefer list layouts over the increasing trend of pie layouts.

We evaluated both layouts in a hierarchical structure, which is widely used in real-world applications. Hierarchical layouts are organised in a nested manner, whereby selecting one option reveals a subsequent menu level containing

additional options. This structure mirrors the natural way humans categorise and access information, making it an intuitive method for managing complex interactions [1]. Although both list and pie layouts can be structured hierarchically, such structures present challenges, including the placement of subsequent menus and the potential for overlapping. While various solutions have been presented for hierarchical lists and pie menus, a direct comparison between them, especially for gaze-based interaction, is lacking.

Kim et al. [30] designed and implemented a novel hierarchical pie menu for border-crossing gaze-based interaction with world-referenced placement. Their innovation included visual anchors used as resting points to address the "overshooting" problem observed in border-crossing menus with sub-menus. However, they focused primarily on the design parameters of their pie menu and compared it only with other border-crossing pie menus. Early adaptations of hierarchical list menus, e.g., [40,56], derived their design directly from 2D desktop environments. Although efficient, these designs can lead to overly complex hierarchical structures, especially in VR [20]. Despite these challenges, hierarchical list menus remain a crucial design element for interactive systems; however, they are mainly designed for controller-based interaction, with a lack of hierarchical list menus tailored for gaze-based interaction. Therefore, in our study, we incorporated design elements from Kim et al. [30] to develop hierarchical list menus specifically suited for gaze-based interaction.

2.2 Menu Manipulation

The increased interest in XR has led to an increase in research focused on natural interaction methods that enable users to interact with computer systems intuitively, mirroring real-world interactions [9], which enhances user immersion in virtual environments. As a result, various methods for controlling virtual menus have been developed, including handheld controllers [42,48], hand gestures [20], gaze-based interaction [30,45,53], speech-based interaction [43], and multimodal systems [44,48].

Interaction via controllers is widely considered the default in VR [20]. However, it is important to explore viable alternatives to controllers to minimise the hardware required for system operation. Hands-free solutions can prove advantageous in scenarios where users' hands might be occupied, and they are particularly beneficial in public setups due to their discreet nature. Regardless of the modality, each interaction method needs to incorporate a way of pointing, such as pointing with a virtual ray at menu items, and a way of selection, such as using a button to confirm the choice. Speech-based and gesture-based methods do not require explicit pointing mechanisms, as they provide direct shortcuts to the corresponding menu options; for example, by uttering a keyword or displaying a hand signal, the linked option is selected, thereby enabling fast interaction. However, these approaches require prior system knowledge and training to operate correctly, which is why we decided not to use them. On the other hand, individuals naturally direct their gaze towards objects they want to interact with, making gaze a seamless and intuitive interaction method. However,

gaze-based interaction is susceptible to the Midas Touch problem [24], which can significantly undermine the usability and efficiency of gaze-based interaction by triggering unintended user inputs. Therefore, different approaches to gaze-based interaction have been developed to mitigate this issue. These approaches can be categorised into Gaze-only interaction [15,28,30] and multimodal interaction [34,37,44,47,53], the latter combining gaze with an additional interaction method.

Gaze-Only Interaction. seeks to remove the need for physical input devices by relying solely on the eye trackers integrated into XR devices. Dwell-based interaction triggers a selection when the user's gaze remains on a button for a set amount of time, and is a well-researched interaction method [21,24,39,44,48,54], which makes it a suitable baseline for comparison. Long dwell times have been shown to reduce the rate of false activations, but they also introduce a delay between the user's action and the system's response. A dwell time above one second is often considered long, while short dwell times can be as brief as 280 ms [39]. For hierarchical menu interaction, shorter dwell times are preferable to reduce the overall interaction time. Majaranta et al. [39] introduced adjustable dwell times for typing tasks, allowing users to manually modify their dwell time via a button press; this resulted in better interaction times without an increase in error rate. However, we used a fixed dwell time, similar to Monteiro et al. [43] and Mutasim et al. [45], to simplify the system learning process.

Border-crossing, another Gaze-only method, triggers selection when the pointer enters an item's selection zone, essentially making it a dwell interaction with a 0 ms dwell time; it has been explored as an interaction method with pie menus [2,21,30]. Although fast, it requires precise control, and it suffers from the "overshooting" problem, where rapid eye movements accidentally trigger multiple menu levels simultaneously. Ahn et al. [2] addressed this by dynamically adjusting the position of subsequent menu levels to the next resting eye position. Kim et al. [30] enhanced this approach by using visual anchors as resting points, eliminating position calculations and reducing interaction times and error rates compared to Ahn et al. [2]. Mutasim et al. [44] reported similar findings by showing that border-crossing is a fast and robust gaze-based interaction method. Smooth pursuits have also been used as an advanced Gaze-only menu interaction method; however, they require special menu layouts with moving options to create trackable trajectories in VR scenes [28].

Multimodal Gaze-Based Interaction. uses eye tracking for pointing or pre-selection, while an additional modality is used to confirm the pre-selection. In gaze-head interaction, one cursor is linked to the eyes and another to the head; both cursors must focus on the same element for a selection to be made. This method builds on the natural coordination between the head and eyes [53] and presents an intuitive and easy-to-learn hands-free approach. However, Lediaeva and LaViola [34] found that head-based selection required more time and was less popular than other methods. Other approaches have combined gaze with

hand-based interaction, e.g., [37,38,47]. These methods, however, are generally inferior to controllers due to low hand-tracking accuracy [50] An alternative multimodal method is the gaze-and-button approach, in which the gaze is used for pre-selection, and a physical button press confirms the selection. This can be implemented using handheld controllers [34,48], which are commonly included with most VR devices, or a keyboard [44]. According to Mutasim et al. [44], the gaze-and-button method is one of the faster interaction techniques; however, it tends to suffer from higher error rates due to hand-eye coordination challenges, where users may press the button before reaching the target or leave the target prematurely. Providing visual feedback, such as highlighting the button with a coloured border when the gaze cursor enters the target, can help mitigate these issues. Despite these performance issues, Pfeuffer et al. [48] reported that users were positive about it, making it the second most preferred method in their study comparing five different methods. We opted for a gaze-and-button approach using a controller (i.e., gaze-and-controller) because it does not require users to learn unfamiliar gestures and offers a balance between Gaze-only and Controller-only methods while avoiding issues related to low hand-tracking accuracy.

Table 1 summarises the key aspects of the most relevant publications and highlights the research gap our study aims to address. Previous work has mainly focused on either pie menus with Gaze-only methods (e.g. Kim et al. [30]'s border-crossing approach) or list layouts with controller-based interaction [42]. Lediaeva and LaViola [34] combined both layouts but omitted hierarchical structures, while Mutasim et al. [44] concentrated solely on target selection. Although gaze-based interaction for hierarchical menu navigation has been explored [32], it was not extensive, and a comprehensive comparison remains lacking. Our study addresses this gap by systematically comparing hierarchical pie and list menu layouts across Gaze-only, Controller-only, and multimodal interaction methods.

Table 1. Comparison of menu design layouts and interaction methods between the most relevant publications and our study setup.

	Layout		Hierarchical	Interaction Method			Visual Anchors
	Pie	List		Gaze-only	Controller-only	Multimodal	
Lediaeva and LaViola [34]	✓	✓	✗	✗	✓	✓	✗
Monteiro et al. [42]	✓	✓	✓	✗	✓	✗	✗
Mutasim et al. [44]	✓	✗	✓	✓	✗	✓	✗
Kim et al. [30]	✓	✗	✓	✓	✗	✗	✓
Our Study	✓	✓	✓	✓	✓	✓	✓

3 Design and Implementation

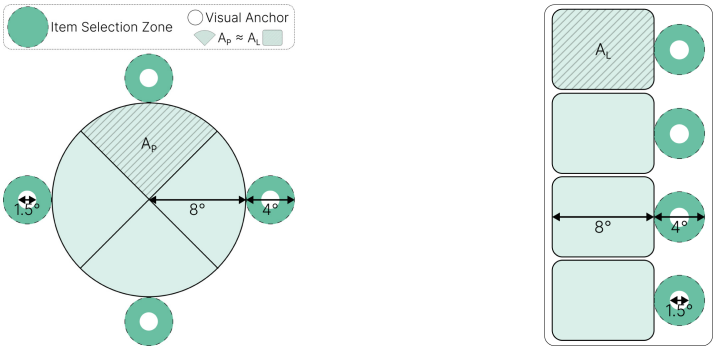
For our implementation, we used the HTC Vive XR Elite headset[3], along with the supplied motion controllers as input devices. The headset was equipped with

[3] https://www.vive.com/eu/product/vive-xr-elite/ (Accessed 17 Feb 2025).

the Vive Facial Tracker[4] for eye tracking capabilities, and it features adjustable diopters, allowing user-specific lens settings for individuals with corrected vision.

3.1 Menu Layout Design

We implemented two distinct menu layouts in a hierarchical structure. The pie menu, shown in Fig. 1a, is based on the lattice menu by Kim et al. [30], featuring a circular arrangement that allows users to select menu options by directing their gaze at visual anchors positioned equidistantly around a central point. The list menu (shown in Fig. 1b) is adapted from Monteiro et al. [42], incorporating visual anchors from Kim et al. [30] within menu options with subsequent levels extending to the right.



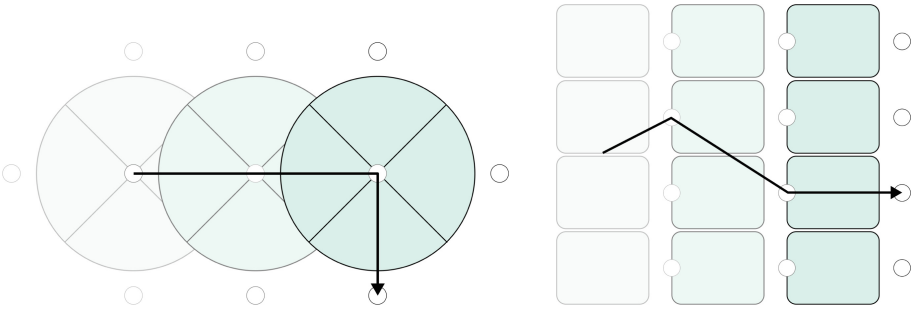
(a) Pie menu (b) List menu

Fig. 1. Design of our (a) pie and (b) list menus, with angles represented in visual degrees. This shows that we kept the dimensions consistent across both layouts.

Following the suggestions of Kim et al. [30], both menu layouts were designed with a horizontal visual angle of 8° for menu items and an additional 4° for the item selection zone in the case of Gaze-only interaction. The separation of the menu item and the item selection zone mitigates accidental menu selections during dwell-based Gaze-only interaction, addressing the Midas Touch problem. Additionally, each item selection zone includes a visual anchor with a radius of 1.5°, serving as a resting point to address the potential "overshooting" problem when navigating hierarchical menus using Gaze-only interaction. Figure 1 provides an overview of the menu layouts, highlighting the item selection zones and visual anchors. Both menus had a total of 12 options distributed across three levels. Figure 2 illustrates how the menus unfold. In both layouts, the subsequent levels unfold, facing the participant to enhance usability and facilitate directional movements [8,20]. List menus, as shown in Fig. 2b, extend to the right, adjacent

[4] https://www.vive.com/eu/accessory/vive-full-face-tracker/ (Accessed 17 Feb 2025).

to the centre of the visual anchor, while the subsequent levels of pie menus, as shown in Fig. 2a, extend in the direction of the selection, positioning themselves at the centre of the visual anchor for easy and intuitive navigation. To facilitate seamless menu navigation and accommodate various interaction types, we incorporated visual feedback into the design. When a user hovers their gaze or controller over a menu option, the option is highlighted, indicating readiness for selection. Additionally, we incorporated a progress bar around the visual anchors, matching the size of the item selection zone for visual feedback during Gaze-only interaction.



(a) Pie menu unfolding (b) List menu unfolding

Fig. 2. The unfolding process of the (a) pie and (b) list menus, showing the gradual expansion of the menu as the items are revealed in sequential steps.

3.2 Interaction Method Design

We investigated three types of interaction methods: (i) Gaze-only dwell-based interaction, (ii) Controller-only interaction, and (iii) a multimodal Gaze-and-Controller interaction, which uses gaze for pointing and a button press from the controller for selection. We initially used border-crossing as an additional Gaze-only interaction method to avoid the Midas Touch problem, but it was only evaluated during the pilot study and not in the main user study.

The Gaze-only dwell-based interaction requires a user to maintain their gaze within the item selection zone for a pre-defined duration before the menu option is selected. We adopted a dwell time of 500 ms, as suggested by Monteiro et al. [43]. The progress bar provides visual feedback to the user during the interaction. This method is well-researched and has a low error rate, making it reliable, though it is inherently slow due to the minimum time required for each selection. Furthermore, for Gaze-only interaction, it is crucial that the menu levels expand, as shown in Fig. 2, and not overlay on top of each other; otherwise, this could lead to multiple unintentional inputs.

The Controller-only interaction is often seen as the most commonly used interaction method in XR. Users select menu options by pointing their controller within the region of a menu option and pressing a button. The visual anchors are removed for Controller-only interaction because the menu options must be selected directly. Similar to Gaze-only interaction, hovered menu items are highlighted to provide visual feedback. Additionally, haptic feedback is provided when the pointer moves over a menu option.

The multimodal Gaze-and-Controller interaction uses gaze as a pointing mechanism and a button press to confirm the selection, addressing the Midas Touch problem and eliminating the need for dedicated item selection zones. Similar to Controller-only interaction, visual anchors are not present because Gaze-and-Controller requires the direct selection of menu items. Additionally, to prevent users from prematurely pressing the button before their gaze reaches the target, similar to what Mutasim et al. [44] reported, visual and haptic feedback are provided when the user looks at a menu option.

3.3 Virtual Environment Design

Our VR scene was developed in Unity 2022.3.13f1[5]. The interaction methods were implemented using the XR Interaction Toolkit v2.5.2[6]. To run the application, we used the Steam VR OpenXR runtime[7], and enabled MR passthrough via Vive Business Streaming[8] to mitigate motion sickness and allow users to see the environment and the controllers. The study setup was powered by a high-performance workstation with an Nvidia RTX 4090, an Intel i9-13900K processor, and 64 GB RAM, which allowed the experimenter to monitor the AR view of the participants and provide guidance when necessary.

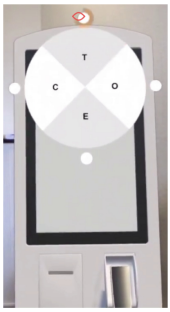
The virtual environment setup, as shown in Fig. 3, included a virtual self-service kiosk displaying instructions, which served as the primary interaction point. The kiosk was positioned 2.2 m from the user, while the interactive menus were placed 1.8 m away. The menus, shown in Fig. 4, were scaled to maintain a uniform size, measured in visual degrees using the following equation: $(\text{visual degree} = 2 \cdot \arctan\left(\frac{\text{size}}{2 \cdot \text{distance}}\right))$, and oriented to face the user, ensuring a consistent experience across all menu levels, as explained in Sect. 3.1. To create a controlled study environment, we disabled locomotion within the virtual space, set a fixed height for the virtual avatar, and conducted the study with participants in a seated position. Additionally, with passthrough enabled, we turned off the virtual rendering of the controllers to prevent any potential confusion and controlled the room lighting for visual consistency. For both Controller-only and Gaze-and-Controller interaction methods, we mapped the primary interaction button from the trigger buttons to the primary "A" or "X" buttons on the right

[5] https://unity.com/releases/editor/whats-new/2022.3.13 (Accessed 17 Feb 2025).
[6] https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@2.5/manual/installation.html (Accessed 17 Feb 2025).
[7] https://store.steampowered.com/steamvr (Accessed 17 Feb 2025).
[8] https://business.vive.com/eu/solutions/streaming/ (Accessed 17 Feb 2025).

or left controller, respectively, to prevent accidental activation from the multiple activation zones of the trigger and support ambidextrous use. In addition to the built-in eye tracker calibration, we implemented a gaze accuracy grid based on [10] to measure the precision and accuracy of the eye trackers after the integrated calibration process of the headset to ensure optimal calibration accuracy. Our source code is publicly available at https://github.com/DFKI-Interactive-Machine-Learning/menu-navigation-in-VR.



Fig. 3. The virtual environment setup for our user study. The scene includes a virtual kiosk at 2.2 m and interactive menus at 1.8 m from the user.



(a) Pie menu



(b) List menu

Fig. 4. Screenshots of the implemented menu designs. (a) shows the pie menu with dwell interaction, featuring the gaze indicator and the radial progress bar; while (b) displays the list menu with controller-based interaction, highlighting a menu option for selection.

4 User Study

Our study goal was to investigate gaze-based menu manipulation in XR setups. The main within-subject user study consisted of six conditions: (i) Pie menu & Gaze-only interaction, (ii) Pie menu & Controller-only interaction, (iii) Pie menu & Gaze-and-Controller multimodal interaction, (iv) List menu & Gaze-only interaction, (v) List menu & Controller-only interaction, and (vi) List menu & Gaze-and-Controller multimodal interaction. The experimental setup, based on Kim et al. [30], evaluated usability, efficiency, and user preferences across these conditions, as described in Sect. 3. The order of the $Layout \times Interaction$ conditions was counterbalanced using a Balanced Latin Square, resulting in six unique sequences, and participants were randomly assigned to one of these sequences to minimise learning and fatigue effects, similar to [30,44].

We initially conducted a within-subject pilot user study with four experienced users to obtain feedback and improve our setup prior to the main study.

The pilot study included eight conditions (i.e., two menu layouts and four interaction methods) due to the inclusion of border-crossing as an additional Gaze-only interaction method. However, based on the results and user feedback, we shortened the study by (i) reducing the number of tasks from ten to five, and (ii) removing the border-crossing interaction method to mitigate user fatigue, as the planned study exceeded the expected 60–80 min. Additionally, (iii) we introduced a short break after each condition, allowing users to remove the headset, which meant that every condition started with eye-tracking calibration and testing. After integrating these modifications, we implemented the main user study. Prior to conducting either study, we obtained approval from our institute's ethical review board. An overview of our study procedure can be seen in Fig. 5.

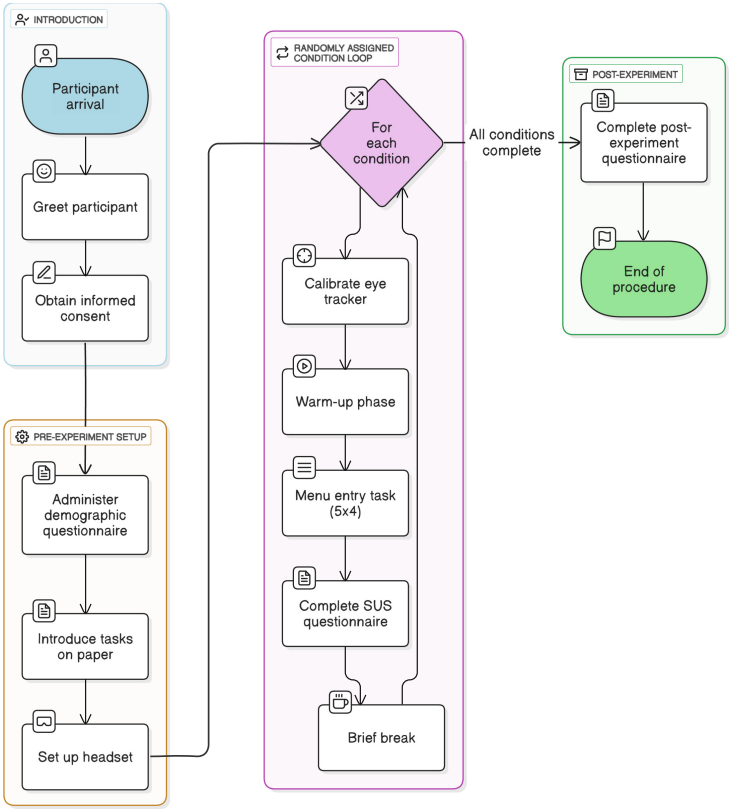


Fig. 5. This is a flowchart depicting our study procedure.

4.1 Hypotheses

We had three main assumptions prior to conducting the user study:

H1 Layout: There is no significant difference in usability and efficiency between pie and list menus, despite the prevalence of pie menus with gaze-based interaction methods in the literature [34,42].
H2 Modality: Incorporating visual and haptic feedback into the multimodal Gaze-and-Controller interaction method will (i) reduce task completion time compared to the Gaze-only dwell-based interaction [44]; (ii) achieve an error rate equivalent to or lower than that of the Gaze-only dwell-based interaction; (iii) enhance usability relative to both Controller-only and Gaze-only methods.
H3 Combination: List menus, being more familiar to users, combined with Gaze-and-Controller interaction, will exhibit the highest usability and user preference among all six conditions.

4.2 Metrics

To evaluate our hypotheses, we collected demographic data and prior XR and eye tracking experience, and measured task completion time, error rate, and System Usability Scale (SUS) scores. The SUS questionnaire consists of ten questions that rate user satisfaction and usability on a 5-point Likert scale [6], and was integrated into the AR environment. We also conducted a post-experiment questionnaire to compare user preferences[9].

4.3 Tasks

Following Kim et al. [30], the menu designs were validated using world-referenced placement, where menus were positioned at fixed coordinates in the AR environment. Each of the six conditions was evaluated using five menu entry tasks, repeated four times. For each menu entry task, participants were given five random three-letter strings, with each letter representing a menu option that can be selected using the respective interaction method. For example, for a string "cat", the first level displays "z, c, t, g" and participants should select "c"; the second level displays "a, b, c, d" and they select "a"; and the third level displays "t, u, v, w" and they select "t". This process was repeated four times for each string, following Kim et al. [30]. The repetition simulated expert user performance, enabling participants to follow familiar trajectories to target items. This allowed us to assess the learning curve over time and evaluate how efficiency changes for each combination during expert use.

4.4 Procedure

The study began by greeting the participants and obtaining their informed consent after fully explaining the goal of the experiment. This was followed by the administration of a demographic questionnaire. The participants were then

[9] We also recorded the gaze coordinates during the menu entries for further insights but did not use them in this study.

introduced to the tasks on paper and familiarised with all six conditions, consisting of the two menu layouts and three interaction methods used in the study. Next, the headset was set up, including adjustments for comfort, dioptre, and eye tracker calibration. The order of the conditions was counterbalanced across the subjects. Prior to each condition, the eye tracker was recalibrated, and the accuracy and precision of the gaze signal were verified using a gaze accuracy grid. Participants then underwent a warm-up phase to familiarise themselves with the menu layout and interaction method before beginning each condition, where each condition comprised five menu entry tasks as described in Sect. 4.3. After completing each condition, the participants filled out a SUS questionnaire. Brief breaks were provided between conditions to mitigate fatigue. At the conclusion of the experiment, participants completed a post-experiment questionnaire to provide feedback on their experience and preferences. All participants' data were pseudo-anonymised by assigning unique identification codes to ensure confidentiality.

4.5 Participants

To compute the required sample size for the user study, we conducted an a priori analysis for six conditions in a within-subjects design using G*Power [16] for a medium effect (Cohen's $d = 0.25$) according to conventions, with an alpha of 0.05, a test power of 0.80, a correlation among repeated measures of 0.5, and a non-sphericity correction (ϵ) of 1, which indicated that 19 participants should be sufficient, this aligns with similar studies [34,44]. We collected data from 20 participants, but we excluded one because they did not perform all six conditions. The remaining 19 participants (13 male, 6 female) were between 18 and 33 years old (μ=25.47). As only 18 users fit a perfect Latin square, additional users were randomly assigned to a sequence. All participants had normal or corrected vision, with 12 participants (63%) wearing glasses. Twelve participants (63%) indicated previous experience with XR; among them, nine participants (75% of the XR-experienced group, or 47% of all participants) also indicated previous eye tracking experience; the high number of experienced participants was due to a similar concurrent campus study. The study took around 90 min, and each participant was compensated for their time with 15€.

5 Results

We collected task completion times, error rates, and SUS scores from all 19 participants across the six different conditions: (i) Pie & Gaze-only, (ii) Pie & Controller-only, (iii) Pie & Gaze-and-Controller, (iv) List & Gaze-only, (v) List & Controller-only, and (vi) List & Gaze-and-Controller. We computed the median task completion time, mean error rate, and mean SUS score for each condition. Additionally, we computed the same metrics for their combined representations, similar to Monteiro et al. [42], to compare the two main aspects, i.e., menu layouts (Pie vs List) and interaction methods (Gaze-only vs Controller-only vs

Gaze-and-Controller). During the study, we collected $19 \times (2 \times 3) \times (5 \times 4) = 2280$ total measurements. We aggregated all $(5 \times 4) = 20$ menu entry tasks by grouping the data by the 19 participants and all six conditions (i.e., the two menu layouts and three interaction methods), resulting in $19 \times (2 \times 3) = 114$ observations. Based on our investigation of the layouts and interactions, we further aggregated the data by taking the mean of the respective conditions.

For the SUS scores, we used two-way repeated measures Analysis of Variance (ANOVA), paired samples t-test[10] for the post-hoc analysis, and Cohen's d [11] for the effect size. Although the SUS distributions appeared slightly skewed, repeated measures ANOVA is robust to moderate deviations from normality when sample sizes are balanced and sphericity is met [5]. Due to the non-normal distribution of the task completion times and error rates, we applied an Aligned Rank Transform (ART) [58] before conducting a two-way repeated measures ANOVA, Wilcoxon signed-rank test [57] for the post-hoc analysis, and the matched pairs rank biserial correlation [27] for the effect size as non-parametric alternatives. In addition, we used Holm's sequential Bonferroni adjustment to correct the p-values for type I errors [18] similar to Lediaeva and LaViola [34]. All of our statistical analyses were conducted using ARTool [26] and the Pingouin library[11].

5.1 Menu Layouts

For menu layouts, the *task completion time* ANOVA($F(1,18)=1.961$, $p<.001$, $\eta_p^2=.034$), and the *error rate* ANOVA($F(1,18)=6.092$, $p<.05$, $\eta_p^2=.044$) showed statistically significant differences with an overall trend towards fewer errors and faster performance with the Pie layout over List. However, the *SUS scores* ANOVA($F(1,18)=.080$, $p=.781$, $\eta_p^2=.001$) showed no statistically significant differences.

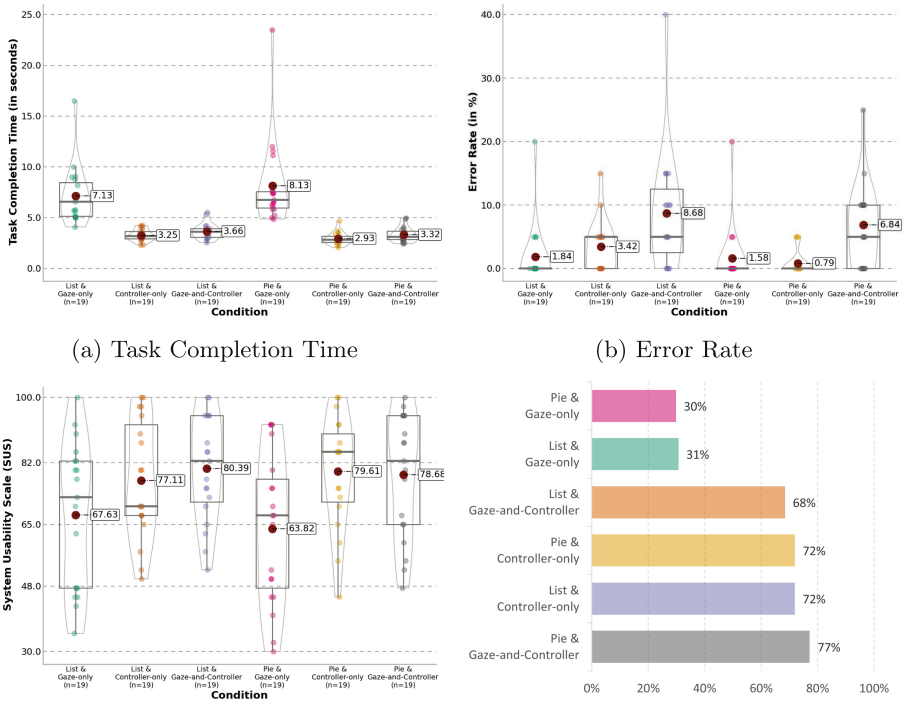
5.2 Interaction Methods

The *task completion time* for the interaction methods produced statistically significant differences ANOVA($F(2,36)=6.444$, $p<.005$, $\eta_p^2=.082$). The post-hoc pair-wise analysis revealed that all three were statistically significant from one another, with Gaze-only being slower than both Gaze-and-Controller Wilcoxon($p<.001$, $r=1.0$) and Controller-only Wilcoxon($p<.001$, $r=1.0$), and Gaze-and-Controller being slower than Controller-only Wilcoxon($p<.001$, $r=.625$).

The results for the *error rates* ANOVA($F(2,36)=.578$, $p=.566$, $\eta_p^2=.008$) showed no statistically significant differences between the interaction methods. However, the *SUS scores* showed statistical significance ANOVA($F(2,36)=7.287$, $p<.005$, $\eta_p^2=.119$), with the post-hoc analysis revealing that the significance was

[10] https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ttest_rel.html (Accessed 17 Feb 2025).

[11] https://pingouin-stats.org/build/html/index.html (Accessed 17 Feb 2025).



(a) Task Completion Time

(b) Error Rate

(c) SUS Score

(d) User Preference

Fig. 6. Results for (a) task completion time, (b) error rate, (c) usability, and (d) user preference across the six conditions. In subfigures (a), (b), and (c), red points indicate mean values. Subfigure (d) shows the ranked order of participants' preferred methods. (Color figure online)

due to Gaze-only having a lower SUS score compared to Controller-only T-test(t=-3.288, p<.005, d=.698)

5.3 Menu and Interaction Combinations

Figure 6 shows the distribution of results from the three evaluation metrics across the six conditions. The *task completion time* (Fig. 6a) showed statistically significant differences ANOVA(F(2,36)=3.925, p<.05, η_p^2=.025) between all six conditions. The post-hoc analysis revealed that List & Gaze-and-Controller was significantly slower than Pie & Gaze-and-Controller Wilcoxon(p<.005, r=.737), and List & Controller-only was significantly slower than Pie & Controller-only Wilcoxon(p<.001, r=.905), but List & Gaze-only and Pie & Gaze-only did not show any statistically significant differences Wilcoxon(p=.418, r=–.226).

However, the *error rates* (Fig. 6b) did not show any statistically significant differences ANOVA(F(2,36)=1.060, p=.357, η_p^2=.029). While for the *SUS scores* (Fig. 6c), we saw that List & Gaze-and-Controller had the highest average

score (μ=80.39, σ=14.56), and Pie & Gaze-only had the lowest average score (μ=63.82, σ=20.60). We can see in Fig. 6c that there are slight differences in the data distributions; however, the differences were not statistically significant ANOVA(F(2,36)=.308, p=.737, η_p^2=.006).

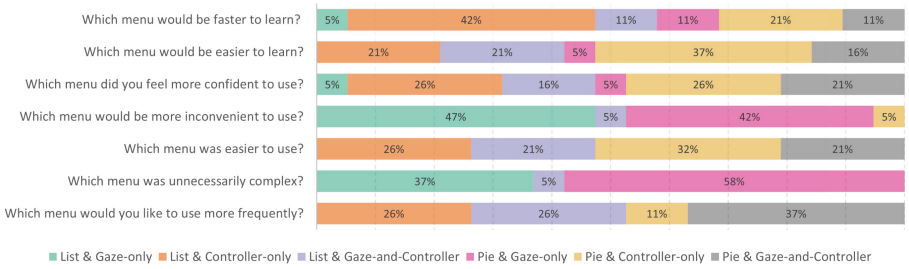


Fig. 7. Participant preferences.

Despite the lack of statistical significance in the usability of the six conditions, noticeable participant preferences emerged in the post-experiment questionnaire (see Figs. 7 and 6d). The responses indicated a clear dislike for both dwell-based Gaze-only menus, with none of the participants preferring to use either of them frequently. All users found them difficult to use, and most (N = 17) considered them inconvenient. However, despite the higher error rates, most participants preferred the Gaze-and-Controller-based combinations (i.e., Pie & Gaze-and-Controller (N = 7) and List & Gaze-and-Controller (N = 5)) for frequent use; this was followed by Controller-only-based combinations (i.e., List & Controller-only (N = 5) and Pie & Controller-only (N = 2)). Regarding ease of use, the preferences were similar: most participants favoured the Pie & Gaze-and-Controller (N = 6), followed by List & Controller-only (N=5), List & Gaze-and-Controller (N = 4), and Pie & Controller-only (N = 4). The post-experiment questionnaire findings can be summarised by the participants' responses to the last point phrased as follows *"Please sort the menus from top (most favourite) to bottom (least favourite)"*. To analyse user preference among the six conditions, we used a weighted rank-order scoring approach. Participants ranked the conditions from most to least preferred, with ranks assigned scores from 6 (most preferred) to 1 (least preferred); we then summed the scores for each condition and normalised these totals by dividing by the maximum possible score, yielding a Normalised Preference Score (NPS) for each condition. This scoring system provides an interpretable measure of relative preference intensity, with higher scores indicating stronger preference. It is evident from the final order shown in Fig. 6d that Pie & Gaze-and-Controller was the most preferred option, while both Gaze-only combinations ranked the lowest.

To evaluate the influence of prior XR and eye tracking experience, we computed point-biserial correlations between these binary factors and both

error rates and task completion times. These analyses revealed no correlation. Responses to the post-experiment questionnaire indicated a slight, non-significant preference for the Gaze-and-Controller interaction among participants without prior eye tracking experience (NPS of 76% vs 69%), while those with such experience displayed a marginally greater preference for the Gaze-only interaction (NPS of 34% vs 27%). Furthermore, participants with previous XR and eye tracking experience favoured Pie menu layouts (NPS of 64%) over List layouts (NPS of 53%), whereas participants without such experience preferred List layouts (NPS of 61%) to Pie layouts (NPS of 56%).

6 Discussion

In this study, we conducted a 19-participant user study to evaluate six conditions formed by two hierarchical menu layouts, and three interaction methods to address our three hypotheses 4.1. Contrary to our initial assumption, Pie menus were statistically faster than List menus for both Controller-only and Gaze-and-Controller interaction methods. In addition, Pie menus were overall statistically less erroneous. This could explain the prevalent use of Pie menus in the literature [2,30,53] despite the reported user preference for List menus [34,42]. Therefore, we reject our hypothesis H1 that both layouts are equivalent.

The Gaze-and-Controller interaction method was significantly faster than Gaze-only. This contrasts with Mutasim et al. [44], who reported no significant speed difference between these methods; it is worth noting that Mutasim et al. [44] used a shorter dwell duration (300 ms vs 500 ms) and focused on a slightly different task (i.e., target selection). However, despite incorporating visual and haptic feedback, the Gaze-and-Controller method exhibited a higher error rate, but it was not statistically significant. When assessing the learning curve, we observed an average reduction in task completion time of approximately 20% across all interaction methods with repetitive menu entries, which aligns with the findings of Kim et al. [30]. Lastly, the post-experiment questionnaire revealed a user preference for the Gaze-and-Controller modality, with the List & Gaze-and-Controller combination achieving the highest SUS scores, thereby supporting hypothesis H3. Therefore, hypothesis H2 cannot be fully retained; although further research is needed to reduce the error rate, the multimodal approach appears to be the most preferred and convenient option for users.

Additionally, Controller-only interaction was significantly faster than both Gaze-only and Gaze-and-Controller, with the Pie & Controller-only condition being the fastest, even outperforming the List & Controller-only condition. This result differs from Monteiro et al. [42], who observed that List menus performed better than Pie menus with Controller-only interaction; this discrepancy might have been caused by our use of unfolding hierarchical levels, as opposed to their overlapping levels, which were unsuitable for Gaze-only.

6.1 Design Guidelines

Our findings indicate that Pie menus are significantly faster with lower error rates compared to List menus. However, for Gaze-only (dwell-based) interaction, there was no statistically significant speed difference between Pie and List layouts. Moreover, users with prior XR/eye tracking experience showed a preference for Pie menus, whereas non-experts tended (non-significantly) to favour List menus. Regarding interaction methods, both the Gaze-only and Controller-only methods achieve comparable accuracy. Controller-only and Gaze-and-Controller are preferable due to their higher speed and usability. Gaze-and-Controller also emerged as the most preferred method among non-experts, while Gaze-only ranked among the least preferred.

Therefore, Pie menus should be used for tasks requiring performance and precision. For Gaze-only (dwell-based) interaction, the layout can be chosen according to other design criteria such as usability or aesthetic considerations, since performance and error rates do not significantly differ. When users are likely to be XR or eye tracking experts, Pie layouts are preferred, but for non-experts, consider using List layouts to align with their intuitive expectations.

Gaze-only interaction should not be used as the primary selection method due to its comparatively lower speed and user preference. Controller-only offers high speed and usability. Alternatively, Gaze-and-Controller offers almost similar efficiency and usability, with a broader user preference. To optimise performance, Pie menus should be paired with Controller-only or Gaze-and-Controller. For novice-oriented interfaces that utilise Gaze-and-Controller interaction, a List layout remains acceptable and aligns with user preference. Additionally, when hands are occupied, or there is motor impairment, Gaze-only, or preferably the multimodal Gaze-and-Controller, should be used since it can be implemented with a single button instead of the controller [44].

Despite the evaluated gaze-based solutions showing lower overall performance, their advantages in terms of privacy, especially in public settings where observation attacks are a concern, are well reported [4,29,31,55]. Our results indicate that using gaze in combination with other modalities is preferred over traditional Controller-only interaction. Although the ethical concerns regarding eye tracking have been discussed in the literature, e.g., [41], our study had minimal impact in this regard and received approval from our institution's ethical review board.

While haptic feedback via controllers and adjustable diopters in the headset may assist users with corrected vision, XR systems, in general, are not optimised for visually impaired users [12,35]. In such cases, a controller should be used, and Gaze-only interaction methods should be avoided due to their lower usability and preference. Furthermore, clear visual feedback should be provided across all interaction methods to reduce errors and enhance task efficiency [33]. Additionally, based on Kim et al. [30], visual anchors should be used to separate item selection zones from the target for Gaze-only interactions; this provides Gaze-only with a comparable error rate to Controller-only.

6.2 Future Work

For Gaze-only interaction, shorter durations, e.g. 300 ms [45], or adjustable dwell times, e.g. [39], may enhance performance and user preference, while incorporating visual indicators could help maintain accuracy [30]. Regarding Gaze-and-Controller, further investigation is needed to reduce error rate, potentially by adding delayed selection or more prominent visual feedback, such as a gaze indicator, to prevent users from looking away when pressing the button.

Our initial removal of the border-crossing-based Gaze-only method was based on its longer task completion times and lower user preference during the pilot study. Our observations indicate that border-crossing was hindered by low eye tracker accuracy, which caused users to spend more time on selection despite conducting a gaze accuracy test prior to each condition. Although our headset's advertised eye-tracker accuracy is higher (0.5°-1.1°)[12] than the device used by Kim et al. [30], users reported low perceived eye-tracking accuracy during both dwell and border-crossing interactions in the pilot study. Enlarging the item selection zone, as suggested by Kim et al. [30], could mitigate this issue, as larger target areas and shorter distances improve ease and speed of selection [8,17]. Future research should evaluate the spatial accuracy and precision of XR headsets, similar to Kapp et al. [25], or adaptively adjust the size of item selection zones based on measured eye tracker accuracy, as suggested by Barz et al. [3].

7 Conclusion

This paper presented a comparative study of gaze-based menu navigation methods in XR, analysing the efficiency, usability, and user preferences of two menu layouts (Pie and List) with three interaction methods: dwell-based Gaze-only, Controller-only, and a multimodal Gaze-and-Controller approach. Through a within-subject study with 19 participants, we conducted statistical analysis to identify key insights and formulate design guidelines for gaze-based menu systems. Our findings indicate that, despite the performance and accuracy of Pie menus, users without prior XR and eye tracking experience favoured List menus. This suggests that users' pre-existing mental models and familiarity with conventional List layouts can outweigh the raw performance benefits of Pie menus, motivating further research toward guided onboarding and adaptive menu systems. While dwell-based interaction demonstrated high accuracy, it received low usability scores and user preference, pointing to the need for further exploration of alternative, innovative interaction methods. Additionally, although controller-based interaction was faster and more accurate than the multimodal approach, users expressed a preference for the gaze-and-controller interaction, highlighting its potential as a preferred discrete selection method in public setups. Overall, our findings contribute to providing insights that can inform future research and development of more effective, user-centred interaction techniques for gaze-based menu systems in XR.

[12] https://www.vive.com/eu/accessory/vive-full-face-tracker/ (Accessed 17 Feb 2025).

Acknowledgments. This work was funded by the European Union under grant number 101093079 (MASTER) and the German Federal Ministry of Education and Research (BMBF) under grant numbers 01IW23002 (No-IDLE) and 01IW24006 (NoIDLEChatGPT), as well as by the Endowed Chair of Applied AI at the University of Oldenburg.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ahlström, D., Cockburn, A., Gutwin, C., Irani, P.: Why it's quick to be square: modelling new and existing hierarchical menu designs. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2010, pp. 1371–1380. ACM, New York, NY, USA, April 2010. https://doi.org/10.1145/1753326.1753534
2. Ahn, S., Santosa, S., Parent, M., Wigdor, D., Grossman, T., Giordano, M.: StickyPie: a gaze-based, scale-invariant marking menu optimized for AR/VR. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, ACM, May 2021. https://doi.org/10.1145/3411764.3445297
3. Barz, M., Daiber, F., Sonntag, D., Bulling, A.: Error-aware gaze-based interfaces for robust mobile gaze interaction. In: Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications, ETRA 2018, pp. 1–10. ACM, New York, NY, USA, June 2018. https://doi.org/10.1145/3204493.3204536
4. Bhatti, O.S., Barz, M., Sonntag, D.: EyeLogin - calibration-free authentication method for public displays using eye gaze. In: ACM Symposium on Eye Tracking Research and Applications, ETRA 2021 Short Papers, pp. 1–7. ACM, New York, NY, USA, May 2021. https://doi.org/10.1145/3448018.3458001
5. Blanca, M., Alarcón, R., Arnau, J., Bono, R., Bendayan, R.: Non-normal data: Is ANOVA still a valid option? Psicothema **4**(29), 552–557 (2017). https://doi.org/10.7334/psicothema2016.383
6. Brooke, J.: SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry, p. 6. CRC Press, 1st edition edn. (1996). https://doi.org/10.1201/9781498710411-35
7. Callahan, J., Hopkins, D., Weiser, M., Shneiderman, B.: An empirical comparison of pie vs. linear menus. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 1988, pp. 95–100. ACM, New York, NY, USA, May 1988. https://doi.org/10.1145/57167.57182
8. Cha, Y., Myung, R.: Extended Fitts' law for 3D pointing tasks using 3D target arrangements. Int. J. Ind. Ergon. **43**(4), 350–355 (2013). https://doi.org/10.1016/j.ergon.2013.05.005
9. Chu, M., Begole, B.: Chapter 17 - natural and implicit information-seeking cues in responsive technology. In: Aghajan, H., Delgado, R.L.C., Augusto, J.C. (eds.) Human-Centric Interfaces for Ambient Intelligence, pp. 415–452. Academic Press, Oxford, January 2010. https://doi.org/10.1016/B978-0-12-374708-2.00017-6
10. Clemotte, A., Velasco, M., Torricelli, D., Raya, R., Ceres, R.: Accuracy and precision of the Tobii X2-30 eye-tracking under non ideal conditions. In: Proceedings of the 2nd International Congress on Neurotechnology, Electronics and Informatics - NEUROTECHNIX, pp. 111–116. SciTePress (2014). https://doi.org/10.5220/0005094201110116, backup Publisher: INSTICC

11. Cohen, J.: Statistical Power Analysis for the Behavioral Sciences. Routledge, New York, 2 edn. July 1988. https://doi.org/10.4324/9780203771587
12. Creed, C., Al-Kalbani, M., Theil, A., Sarcar, S., Williams, I.: Inclusive AR/VR: accessibility barriers for immersive technologies. Univ. Access Inf. Soc. **23**(1), 59–73 (2024). https://doi.org/10.1007/s10209-023-00969-0
13. Duchowski, A.T.: A breadth-first survey of eye-tracking applications. Behav. Res. Methods Instrum. Comput. **34**(4), 455–470 (2002). https://doi.org/10.3758/BF03195475
14. Duchowski, A.T.: Gaze-based interaction: a 30 year retrospective. Comput. Graph. **73**, 59–69 (2018). https://doi.org/10.1016/j.cag.2018.04.002
15. Elmadjian, C., Morimoto, C.H.: Gazebar: Exploiting the midas touch in gaze interaction. In: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, CHI 2021, ACM, May 2021. https://doi.org/10.1145/3411763.3451703
16. Faul, F., Erdfelder, E., Buchner, A., Lang, A.G.: Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav. Res. Methods **41**(4), 1149–1160 (2009). https://doi.org/10.3758/BRM.41.4.1149
17. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. J. Exp. Psychol. **47**(6), 381–391 (1954). https://doi.org/10.1037/h0055392, place: US Publisher: American Psychological Association
18. Holm, S.: A simple sequentially rejective multiple test procedure. Scandinavian J. Stat. **6**(2), 65–70 (1979), publisher: [Board of the Foundation of the Scandinavian Journal of Statistics, Wiley]
19. Holmqvist, K., Nystrom, M., Andersson, R., Dewhurst, R., Jarodzka, H., Weijer, van de, J.: Eye Tracking: A Comprehensive Guide to Methods and Measures. Oxford University Press, Oxford, New York, November 2011
20. Hou, S., Thomas, B.H., Lu, X.: VRMenuDesigner: a toolkit for automatically generating and modifying VR menus. In: 2021 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR), pp. 154–159, November 2021. https://doi.org/10.1109/AIVR52153.2021.00036
21. Huckauf, A., Urbina, M.H.: Gazing with pEYEs. In: Proceedings of the 2008 symposium on Eye tracking research & applications - ETRA 2008, ACM Press (2008). https://doi.org/10.1145/1344471.1344483
22. Imamura, S., Jieun, L., Rekimoto, J., Makoto, I.: Advantage of gaze-only content browsing in VR using cumulative dwell time compared to hand controller. In: Proceedings of the 2023 ACM Symposium on Spatial User Interaction, SUI 2023, pp. 1–8. ACM, New York, NY, USA, October 2023. https://doi.org/10.1145/3607822.3614513
23. Isomoto, T., Yamanaka, S., Shizuki, B.: Interaction design of dwell selection toward gaze-based AR/VR interaction. In: 2022 Symposium on Eye Tracking Research and Applications, ETRA 2022, pp. 1–2. ACM, New York, NY, USA, June 2022. https://doi.org/10.1145/3517031.3531628
24. Jacob, R.J.K.: 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 1990, pp. 11–18. ACM, New York, NY, USA, March 1990. https://doi.org/10.1145/97243.97246
25. Kapp, S., Barz, M., Mukhametov, S., Sonntag, D., Kuhn, J.: ARETT: augmented reality eye tracking toolkit for head mounted displays. Sensors **21**(6), 2234 (2021). https://doi.org/10.3390/s21062234, number: 6 Publisher: Multidisciplinary Digital Publishing Institute

26. Kay, M., Elkin, L., Higgins, J.J., Wobbrock, J.O.: mjskay/ARTool: ARTool 0.11.2, April 2025. https://doi.org/10.5281/ZENODO.594511, https://zenodo.org/doi/10.5281/zenodo.594511
27. Kerby, D.S.: The simple difference formula: an approach to teaching nonparametric correlation. Compr. Psychol. **3**, 11.IT.3.1 (2014). https://doi.org/10.2466/11.IT.3.1, publisher: SAGE Publications Inc
28. Khamis, M., Oechsner, C., Alt, F., Bulling, A.: Vrpursuits: interaction in virtual reality using smooth pursuit eye movements. In: Proceedings of the 2018 International Conference on Advanced Visual Interfaces, AVI 2018, ACM, May 2018. https://doi.org/10.1145/3206505.3206522
29. Khamis, M., et al.: CueAuth: comparing touch, mid-air gestures, and gaze for cue-based authentication on situated displays. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. **2**(4), 174:1–174:22 (2018). https://doi.org/10.1145/3287052
30. Kim, T., Ham, A., Ahn, S., Lee, G.: Lattice menu: a low-error gaze-based marking menu utilizing target-assisted gaze gestures on a lattice of visual anchors. In: CHI Conference on Human Factors in Computing Systems, ACM, April 2022. https://doi.org/10.1145/3491102.3501977
31. Kopácsi, L., Schneider, T.S., Karr, C., Barz, M., Sonntag, D.: Gazelock: gaze- and lock pattern-based authentication. In: Proceedings of the 30th ACM Symposium on Virtual Reality Software and Technology. VRST 2024, ACM, New York, NY, USA (2024). https://doi.org/10.1145/3641825.3689520
32. Kopácsi, L., Klimenko, A., Barz, M., Sonntag, D.: Exploring gaze-based menu navigation in virtual environments. In: Proceedings of the 2024 ACM Symposium on Spatial User Interaction, SUI 2024, pp. 1–2. ACM, New York, NY, USA, October 2024. https://doi.org/10.1145/3677386.3688887
33. Lankes, M., Ramirez Gomez, A.: GazeCues: exploring the effects of gaze-based visual cues in virtual reality exploration games. Proc. ACM Hum.-Comput. Interact. **6**(CHI PLAY), 237:1–237:25 (2022). https://doi.org/10.1145/3549500
34. Lediaeva, I., LaViola, J.: Evaluation of body-referenced graphical menus in virtual environments. In: Proceedings of Graphics Interface 2020, GI 2020, pp. 308 – 316. Canadian Human-Computer Communications Society/Société canadienne du dialogue humain-machine (2020). https://doi.org/10.20380/GI2020.31
35. Liu, T., Fazli, P., Jeong, H.: Artificial intelligence in virtual reality for blind and low vision individuals: literature review. Proc. Hum. Factors Ergon. Soc. Ann. Meeting **68**(1), 1333–1338 (2024). https://doi.org/10.1177/10711813241266832
36. Lu, F., Davari, S., Bowman, D.: Exploration of techniques for rapid activation of glanceable information in head-worn augmented reality. In: Symposium on Spatial User Interaction. SUI 2021, ACM, November 2021. https://doi.org/10.1145/3485279.3485286
37. Lystbæk, M.N., Rosenberg, P., Pfeuffer, K., Grønbæk, J.E., Gellersen, H.: Gaze-hand alignment: combining eye gaze and mid-air pointing for interacting with menus in augmented reality. Proc. ACM Hum.-Comput. Interact. **6**(ETRA), 1–18 (2022). https://doi.org/10.1145/3530886
38. Lystbæk, M.N., et al.: Hands-on, hands-off: gaze-assisted bimanual 3d interaction. In: Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology, pp. 1–12. ACM, Pittsburgh PA USA, October 2024. https://doi.org/10.1145/3654777.3676331
39. Majaranta, P., Ahola, U.K., Špakov, O.: Fast gaze typing with an adjustable dwell time. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2009, ACM, April 2009. https://doi.org/10.1145/1518701.1518758

40. Mine, M.R.: ISAAC: a virtual environment tool for the interactive construction of virtual worlds. Technical Report, University of North Carolina at Chapel Hill, USA, April 1995
41. Mohamed Selim, A., Barz, M., Bhatti, O.S., Alam, H.M.T., Sonntag, D.: A review of machine learning in scanpath analysis for passive gaze-based interaction. Front. Artif. Intell. **7** (2024). https://doi.org/10.3389/frai.2024.1391745, publisher: Frontiers
42. Monteiro, P., Coelho, H., Goncalves, G., Melo, M., Bessa, M.: Comparison of radial and panel menus in virtual reality. IEEE Access **7**, 116370–116379 (2019). https://doi.org/10.1109/access.2019.2933055
43. Monteiro, P., Gonçalves, G., Peixoto, B., Melo, M., Bessa, M.: Evaluation of hands-free vr interaction methods during a fitts' task: efficiency and effectiveness. IEEE Access **11**, 70898–70911 (2023). https://doi.org/10.1109/access.2023.3293057
44. Mutasim, A., Batmaz, A.U., Mughrabi, M.H., Stuerzlinger, W.: Performance analysis of saccades for primary and confirmatory target selection. In: 28th ACM Symposium on Virtual Reality Software and Technology, ACM, November 2022. https://doi.org/10.1145/3562939.3565619
45. Mutasim, A.K., Batmaz, A.U., Stuerzlinger, W.: Pinch, click, or dwell: comparing different selection techniques for eye-gaze-based pointing in virtual reality. In: ACM Symposium on Eye Tracking Research and Applications, ACM, May 2021. https://doi.org/10.1145/3448018.3457998
46. Paulus, Y.T., Remijn, G.B.: Usability of various dwell times for eye-gaze-based object selection with eye tracking. Displays **67**, 101997 (2021). https://doi.org/10.1016/j.displa.2021.101997
47. Pfeuffer, K., Mayer, B., Mardanbegi, D., Gellersen, H.: Gaze + pinch interaction in virtual reality. In: Proceedings of the 5th Symposium on Spatial User Interaction, SUI 2017, ACM, October 2017. https://doi.org/10.1145/3131277.3132180
48. Pfeuffer, K., Mecke, L., Delgado Rodriguez, S., Hassib, M., Maier, H., Alt, F.: Empirical evaluation of gaze-enhanced menus in virtual reality. In: 26th ACM Symposium on Virtual Reality Software and Technology, pp. 1–11. ACM, Virtual Event Canada, November 2020. https://doi.org/10.1145/3385956.3418962
49. Plopski, A., Hirzle, T., Norouzi, N., Qian, L., Bruder, G., Langlotz, T.: The eye in extended reality: a survey on gaze interaction and eye tracking in head-worn extended reality. ACM Comput. Surv. **55**(3), 53:1–53:39 (2022). https://doi.org/10.1145/3491207
50. Rantamaa, H.R., Kangas, J., Kumar, S.K., Mehtonen, H., Järnstedt, J., Raisamo, R.: Comparison of a vr stylus with a controller, hand tracking, and a mouse for object manipulation and medical marking tasks in virtual reality. Appl. Sci. **13**(4), 2251 (2023). https://doi.org/10.3390/app13042251
51. Rauschnabel, P.A., Felix, R., Hinsch, C., Shahab, H., Alt, F.: What is XR? towards a framework for augmented and virtual reality. Comput. Hum. Behav. **133**, 107289 (2022). https://doi.org/10.1016/j.chb.2022.107289
52. Reiter, K., Pfeuffer, K., Esteves, A., Mittermeier, T., Alt, F.: Look & turn: one-handed and expressive menu interaction by gaze and arm turns in VR. In: 2022 Symposium on Eye Tracking Research and Applications, ACM, June 2022. https://doi.org/10.1145/3517031.3529233
53. Sidenmark, L., Potts, D., Bapisch, B., Gellersen, H.: 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. ACM, May 2021. https://doi.org/10.1145/3411764.3445697

54. Starker, I., Bolt, R.A.: A gaze-responsive self-disclosing display. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 1990, pp. 3–10. ACM, New York, NY, USA, Mar 1990. https://doi.org/10.1145/97243.97245
55. Stephenson, S., Pal, B., Fan, S., Fernandes, E., Zhao, Y., Chatterjee, R.: SoK: authentication in augmented and virtual reality. In: 2022 IEEE Symposium on Security and Privacy (SP), pp. 267–284, May 2022. https://doi.org/10.1109/SP46214.2022.9833742, iSSN: 2375-1207
56. van Teylingen, R., Ribarsky, W., van der Mast, C.: Virtual data visualizer. IEEE Trans. Visual Comput. Graphics **3**(1), 65–74 (1997). https://doi.org/10.1109/2945.582350
57. Wilcoxon, F.: Individual comparisons by ranking methods. Biometrics Bull. **1**(6), 80–83 (1945). https://doi.org/10.2307/3001968, publisher: [International Biometric Society, Wiley]
58. Wobbrock, J.O., Findlater, L., Gergle, D., Higgins, J.J.: The aligned rank transform for nonparametric factorial analyses using only anova procedures. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2011, pp. 143–146. ACM, New York, NY, USA, May 2011. https://doi.org/10.1145/1978942.1978963

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

