Exploring the Impact of Distance on XR Selection Techniques

Becky Spittle
HCI Research Centre, College of Computing
Birmingham City University
Birmingham, United Kingdom
becky.spittle@bcu.ac.uk

Maite Frutos-Pascual
HCI Research Centre, College of Computing
Birmingham City University
Birmingham, United Kingdom
maite.frutos@bcu.ac.uk

Chris Creed
HCI Research Centre, College of Computing
Birmingham City University
Birmingham, United Kingdom
chris.creed@bcu.ac.uk

Ian Williams
HCI Research Centre, College of Computing
Birmingham City University
Birmingham, United Kingdom
ian.williams@bcu.ac.uk

Abstract

Designing eXtended Reality (XR) interaction techniques that function efficiently across varying contexts remains a significant challenge, leading to a fragmented landscape of input paradigms. A key factor influencing interaction performance is the user's distance from virtual content. To address this, we present two user studies investigating how distance impacts interaction efficacy and user experience in XR environments. These studies evaluate widely used interaction methods—freehand techniques (Press, Airtap, Hover) and gaze-based techniques (Eye and Head)—across four proxemic zones: Intimate, Personal, Social, and Public. Each study involved 32 participants, with a combined total of 4,608 interaction trials (1,152 in Study 1 and 3,456 in Study 2). Findings reveal the strengths and limitations of techniques depending on user-object distance. Based on these insights, we offer design recommendations for tailoring XR interaction modalities to proxemic factors. This promotes adaptable interaction design, enabling more inclusive and personalised user experiences across diverse XR scenarios.

CCS Concepts

• **Human-centered computing** → **Mixed / augmented reality**; **Empirical studies in interaction design**; **Interaction design theory, concepts and paradigms**.

Keywords

Augmented Reality, Mixed Reality, Extended Reality, Distance, Proxemic Interaction, Interaction Techniques, Adaptive, Context

ACM Reference Format:
Becky Spittle, Maite Frutos-Pascual, Chris Creed, and Ian Williams. 2025. Exploring the Impact of Distance on XR Selection Techniques. In *Proceedings of the 27th International Conference on Multimodal Interaction (ICMI '25), October 13–17, 2025, Canberra, ACT, Australia.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3716553.3750770

1 Introduction

Recent advances in immersive technologies have rapidly evolved eXtended Reality (XR) from concept to a commercially viable medium, with significant implications for the future of human-computer interaction (HCI) [1, 58]. By merging physical and digital worlds, XR is redefining how users can engage with digital content in everyday contexts. Current XR systems are often equipped to support environmental awareness, motion tracking, and multimodal interaction [13]. However, many XR applications still rely on a single, fixed input method, limiting their ability to respond to users' changing needs and interaction contexts [18, 52].

To unlock XR's potential as a ubiquitous, user-centric technology, interactive systems should become capable of interpreting the spatial relationships between users and surrounding digital or physical elements. This would allow systems to offer the most appropriate input technique or system output, tailored to contextual factors such as user abilities, preferences, or environment [8, 9, 16, 30].

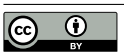
This idea of spatially adaptive interaction design draws heavily from Proxemics [17], a theory rooted in social science that classifies interpersonal space into four zones: Intimate (<0.5 m), Personal (0.5–1 m), Social (1–4 m), and Public (>4 m) [10]. These zones shape how people behave and interact depending on distance, activity, and perceived affordances [15].

In ubiquitous computing, Proxemics has been used to support context-aware interactions that adapt as people and devices move through space [2, 14]. However, its integration into XR interaction design, particularly for switching or combining inputs based on distance, remains under-explored. While studies have considered distance-aware interfaces, focus has primarily been given to output adaptations [25] and freehand paradigms [8], with no studies found to compare a range of commonplace, peripheral-free selection techniques across proxemic zones.

To begin addressing this gap, this paper reports on two empirical studies investigating how distance affects fundamental selection. Conducted in a seated setup, where users cannot easily adjust their position, the studies evaluate freehand (Press, Airtap, Hover) and gaze-based techniques (Eye and Head) across all proxemic zones. Selection time, error rate, task load and preference are used to compare the performance and suitability of each input method. This provides a foundation for exploring how proxemic dimensions can inform context-aware interaction techniques in XR.

2 Background

Distance has been found to affect the suitability of XR input techniques, thus playing a crucial role in selecting appropriate methods for XR interactions in different spatial contexts [24, 63]. In this



This work is licensed under a Creative Commons Attribution 4.0 International License.
ICMI '25, Canberra, ACT, Australia
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1499-3/25/10
https://doi.org/10.1145/3716553.3750770

section, focus is given to highlighting the advantages and limitations of freehand and gaze techniques across different distances and applications.

2.1 Freehand Interaction

Freehand interaction in XR, using gestures such as pressing, pinching, pointing, and grabbing, is widely supported in head-worn displays (HWDs) and allows users to intuitively perform actions like selecting, manipulating, or scaling virtual objects across different distances [18, 23, 62]. These gestures are often described as engaging and immersive, especially in creative or exploratory contexts such as gaming or artistic tasks [19, 47, 55]. Many techniques are based on familiar input metaphors; for instance, the Airtap gesture replicates a mouse click and has been found to be as effective as physical buttons in various conditions [12, 18, 42].

Studies comparing near and far gesture techniques report minimal differences in efficiency or usability. Caputo et al. [7] found that performance was not significantly affected by distance when evaluating freehand methods such as Airtap for room-scale manipulation. Similarly, Lilligreen et al. [26] observed no substantial differences between Press for near targets and Airtap for distant ones, although some participants felt that Airtap was less intuitive. Gestures such as pinching may also lead to unintentional activations when the user's hands are at rest [45].

Simplified techniques like Hover aim to reduce gesture complexity by allowing users to select objects by holding a pointer over them for a short duration [48]. While Hover can be effective for selecting large objects [3], it often suffers from issues such as the Midas touch [4] and can cause fatigue due to prolonged arm elevation [48]. It is also less suitable for tasks that require fine precision [29], indicating that its use may be best limited to specific scenarios.

To improve the overall user experience, researchers have explored adaptive systems that change input methods based on user proximity to virtual content [8]. For example, the HoloLens 2 encourages the use of Press for nearby objects and Airtap for distant ones [41]. In addition, refinement methods like Go-Go allow users to extend their reach through non-linear mappings [56], and dynamic scaling techniques adjust the size or placement of virtual elements according to distance [28, 61]. Since no single gesture performs optimally in every context, future XR systems are likely to benefit from multimodal interaction approaches that adapt input methods to both user preferences and situational needs [16, 30].

2.2 Gaze Interaction

While typically used for implicit input [44], head and eye tracking are increasingly employed for explicit interactions in XR [18]. These techniques are particularly valuable when users' hands are occupied, enabling broader use cases than freehand interaction alone [43, 50]. Gaze-based interaction is discreet and useful for quick selections, such as pressing buttons or navigating menus [29], but is often viewed as less intuitive and harder to implement for complex tasks [32, 50]. Despite these limitations, users generally find gaze input easy to learn [44].

Head tracking is more stable across environments due to the nature of the built-in HWD sensors used, making it more accurate and consistent than hand or eye input for distant targeting [29, 59]. For example, Whitlock et al. [63] explored the impact of distance on techniques, comparing head-based pointing (with Airtap or Voice as a selection mechanism) and a hardware controller, for objects placed in Social and Public proxemic zones. Although handheld remotes are commonplace, they found that participants performed comparably when using head-gaze as a pointing method, and it was perceived as significantly more efficient and usable than device-mediated interactions, primarily due to cursor visibility at the centre of the user's field of view.

Eye tracking, now a common feature in HWDs such as the HoloLens 2 and Apple Vision Pro, often struggles to differentiate between user attention and intention, particularly when using dwell-based selection [4]. Pfeuffer et al. [44] explain that this challenge arises because XR gaze interactions involve two distinct tasks, information consumption and interface selection, requiring careful consideration of the transitions between the real world and the virtual interface. Inaccuracies are also caused by environmental noise, calibration demands, and reduced reliability when targeting small or distant objects [24, 59]. These challenges often lead users to prefer hybrid methods that combine gaze with a confirmation input like blinking or button presses [31].

Empirical comparisons suggest that head pointing generally outperforms eye-only methods, especially for distant interactions [24]. Qian and Teather [49] compared eye, head, and a multimodal technique (using both eye and head) for selections in the Public zone. Results indicated that eye-only selection provided the worst performance in terms of error rate, selection times, and throughput, which was primarily attributed to issues with eye tracking and calibration. The hybrid approach, combining eye and head, failed to leverage the benefits of both methods, with head offering the fastest selection times, the best accuracy, and being strongly preferred.

Although head and eye offer effective pointing techniques, the potential of using these inputs interchangeably remains underexplored. As XR systems evolve, more research is needed to understand how different inputs can be integrated into adaptive interaction frameworks that respond to distance and context [30, 52].

3 Methodology

Two user studies were designed to explore the appropriateness of commonplace interaction techniques at different distances. Seated scenarios are prevalent in XR (e.g., for extended workspaces [31, 35] and interaction while travelling [8]), therefore, focus is given to fundamental selection within a seated MR environment.

In Study 1, objects are selected within the Intimate proxemic zone (<0.5m) using Press and three time-based selection techniques; Hover, Head and Eye. Study 2 compares Airtap, Hover, Head and Eye with objects placed within the Personal (0.5m-1.0m), Social (1.0m-4.0m) and Public (>4.0m) proxemic zones. (see Figure 1).

Two studies were conducted to limit session duration (prioritising participant well-being) and to prevent extra cognitive load. Combining near and far interactions would have forced participants to switch between Airtap and Press, adding complexity that might have biased comparisons at this stage of research.

The task explored in both studies is centred around a widespread "observe and interact" scenario, where a user may be referring to

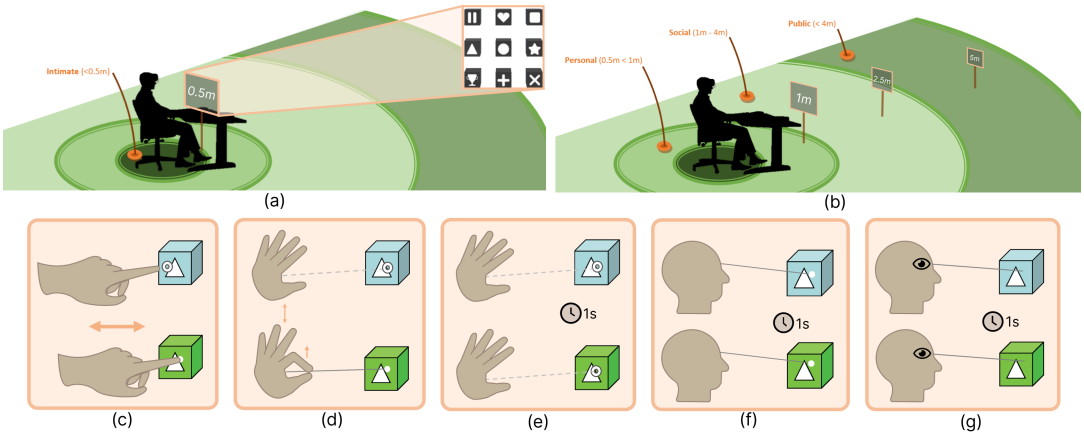


Figure 1: Study 1 within the Intimate proxemic zone (a) evaluates four techniques (Press (c), Hover (e), Head (f), and Eye (g)), while Study 2 across Personal, Social, and Public zones (b) evaluates Airtap (d), Hover (e), Head (f), and Eye (g). Both studies involved selecting 9 cubes at each distance from an array of objects as shown in (a).

one interface component to gather information (e.g. a text document/set of instructions, specification, video, or design concept), before interacting with another to perform a task (e.g. selecting a virtual menu panel/button or an object to interact with) [8, 27]. Details for how both studies were conducted are provided below.

3.1 Apparatus

In both studies, the XR environment was developed for the Microsoft HoloLens 2 using C# in Unity and incorporating the MRTK 2.7.2.0 SDK. Throughout each study session, the Windows Device Portal was used to remotely launch applications, monitor participants, and assess the stability and refresh rate of the HWD. First-person, live Mixed Reality captures were recorded via the device portal to support data analysis. Additionally, ambient lighting was routinely measured with a lux meter to ensure it remained within the recommended range of 500-1000 lux [37].

3.2 Environment

The user experiments were conducted in a controlled environment under lab conditions. The room had no furniture except a partition screen blocking the researcher from the participants view, a small desk, and a non-swivel office chair (see Figure 2).

3.3 Participants

32 right-handed participants (21 male and 11 female), aged 22-44 (M = 29.28, SD = 6.01) took part in Study 1. Participants were from a population of university students and staff from the faculty of computing. Based on self-reported experience, participants ranged from novice XR users who had never (n = 6) or rarely (n = 8) used immersive technologies, to more experienced users who used them monthly (n = 12), weekly (n = 5) or daily (n = 1).

Study 2 again included 32 right-handed participants (22 male and 10 female), aged between 22 and 44 (M = 29.13, SD = 6.03). Participants were from the same population of university students and staff as Study 1, with 26 participants taking part in both studies. Users ranged from novices who had never (n = 5) or rarely (n = 6) used immersive technologies, to more experienced users who used them monthly (n = 15), weekly (n = 5) or daily (n = 1).

Handedness was not part of recruitment criteria, and as the object array was 3x3 and centralised (see Figure 2) it is expected for results to be transferable to left-handed users. No participants with colour blindness and/or with corrected visual acuity below 0.80 were included in the analysis (see section 3.6 for details on screening). All participants were compensated with a £10 gift voucher.

3.4 Interaction Techniques

Figures 1 c, e, f, and g illustrate the four interaction techniques explored in Study 1; with d, e, f and g showing those considered in Study 2. Here, focus is given to comparing the widely used baseline techniques of Press and Airtap with common time-based techniques: Hover (hand pointing), Head gaze and Eye gaze [18].

These techniques were developed following HoloLens 2 guidelines, using MRTK 2.7.2.0 [38]. The MRTK provides assets that facilitate accelerated cross-platform development. Therefore, to enhance replicability, the default parameters offered by the MRTK, including cursor sizes and settings, were used. Dwell time for Hover, Head, and Eye was determined by following HoloLens 2 recommendations [39]. Details for each technique are outlined below:

Hand Press. Required participants to position their finger over the target cube before employing a "Press" gesture to execute the selection. The cursor would transition from a torus to a solid circle to communicate depth, with selections being triggered on release.

Hand Airtap. Involved a two-step interaction process. First, users pointed at the object of interest, then employed a "Pinch" gesture to execute the selection. This was achieved by controlling a cursor, which was attached to the end of a ray rendered from the palm. A solid ray and cursor indicated successful gesture registration, with selections being triggered on release.

Hand Hover. Employed a dwell or "Hover" time, requiring that the default MRTK cursor, which was attached to the end of a ray

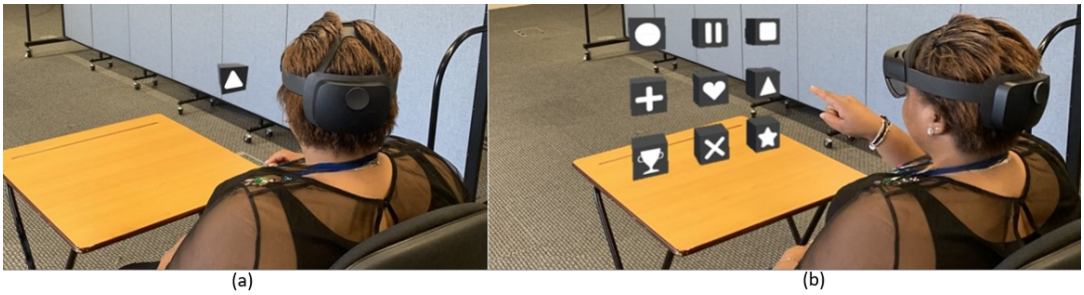


Figure 2: Environment during (a) the observe stage, and (b) interact stage. Objects are representative (not to scale).

rendered from the palm, continuously intersected the object for one second to trigger a selection.

Head Gaze. Employed a cursor attached to the end of an invisible ray, which extended from the HWD towards the forward facing direction. Selections were triggered when the cursor had continuously intersected the object for one second.

Eye Gaze. Relied on the visual feedback defined in Section 3.5 to indicate object targeting and selection. A cursor was attached to the end of a ray which extended from the eyes toward the viewing direction, however, this was not visible to avoid an effect described as "fleeing cursor" [40]. Selections were triggered when participants' gaze continuously intersected the object for one second.

3.5 Tasks

Once comfortably wearing the HoloLens and seated at the desk, participants were asked to look straight ahead at a marker whilst the application loaded. Following this, a text prompt instructed participants to select a start button to begin (using the interaction technique explored) and the environment would alternate between the observe and interact stage as shown in Figure 2.

"Observe" stage. For each trial, participants were instructed to turn to the right (45°) and view the reference stimulus (target), which was displayed for 5 seconds based on BT.500 guidelines [57]. Following this, a text prompt was displayed, instructing the participant to face forward and select the corresponding target from an array of cubes (as considered in previous work [6]).

"Interact" stage. In both studies, the object array measured 25cm x 25cm x 5cm and each cube was spaced 5cm apart. The array was positioned so all cubes were identifiable within the HoloLens 2 FOV, centred -0.12m below the horizon line relative to the height of individual participants [36]. Objects were world-anchored at each distance in accordance with real-world interactions [17, 63].

In Study 1, the array was positioned 0.5m away from the user and cubes had a visual angle of 5.7°. In Study 2, it was placed 1.0m away from the user for interactions within the Personal zone (cubes = 2.9° visual angle), 2.5m away for interactions within the Social zone (cubes = 1.1° visual angle) and 5.0m away from the user for the Public zone (cubes = 0.6° visual angle).

Targets were represented by 9 different symbols (see Figure 2) and were shuffled for each trial. Participants were asked to select objects in randomised order. This was based on object position in Study 1, and object distance and position in Study 2.

Selections could be made by interacting with any point of the target cube. All cubes were black by default and turned blue when targeted, with visual feedback initiated after a delay of 200ms. If a selection was successful, the target cube turned green and a confirmation sound was triggered from the MRTK sound library. In cases where participants interacted with the wrong cube, the target turned red and produced an MRTK error sound.

Participants were asked to return to their natural resting position between trials. A reset task button was also provided at the starting position, which was used in cases where participants were unable to recall the target object. The correct target had to be selected to continue. Participants performed a total of 9 selection tasks with each technique at each distance. After all targets were selected an "End of Condition" message appeared to prompt participants to cease the process.

3.6 Protocol

Each session in Study 1 lasted around 60 minutes, with sessions in Study 2 lasting between 90 and 120 minutes; including rest breaks. The study protocol employed for both studies had received IRB approval and was broken down into the following stages:

Pre-test. Informed consent and demographic information, including participants' prior experience, were collected before each study session. Upon arrival, participants were briefed about the purpose and test protocol of the study. Visual acuity and colour blindness were then assessed using a Snellen chart and the Ishihara test. Following this, participants were seated at a desk and instructed to comfortably adjust and wear the HWD.

Training. After the experimenter explained the technique being tested, participants engaged in a training phase where they were given time to practice and ask questions. The training phase simulated the main experimental task but involved only three selections, and used a different set of symbols to avoid artificial learning effects. Before the first training session, each participant completed eye calibration to optimise the performance of HoloLens 2. Calibration was repeated as necessary.

Test. After attaining verbal confirmation that participants clearly understood the technique and were comfortable with the interaction space, they were presented with the main experimental task (see Figure 2). In Study 1, participants completed 9 selections with

each technique (see Figure 1), which produced 1,152 trials (32 participants x 9 selections x 4 techniques = 1,152). In Study 2, participants completed 9 selections using each technique at 3 distances (1m, 2.5m, 5m). This produced 3,456 trials (32 participants x 9 selections x 3 distances x 4 techniques). After each condition, participants completed a NASA-TLX questionnaire and noted what they liked and disliked about the technique. They would then have a short break before beginning the next condition.

Study 1 followed a within-subjects design, where technique (counterbalanced using a latin square) was defined as the independent variable, and selection time, error rate, task load, and preference were captured as dependent variables. Study 2 followed a within-subjects, mixed-factorial design, where technique and distance were defined as independent variables. Dependent variables were identical to those captured in Study 1.

Post-test. Following the completion of all four conditions, in both studies, a semi-structured interview was conducted to gather subjective feedback and overall preference rankings. This was to better understand interaction approaches and contextualise what participants liked and disliked about each technique.

3.7 Metrics

The following metrics were captured for each technique, capturing factors relating to effectiveness, efficiency and satisfaction [51]:

Task Completion Times. Represent the duration in ms from when each trial began (following the "observe" stage as defined in section 3.5) to when the correct object was selected.

Error Rate. Defines the number of trials where an incorrect selection was made before the correct selection.

Task Load. Measured via the official iOS NASA-TLX (Task Load Index) application, where raw scores were analysed for individual subscales as well as weighted scores for overall task load [22].

Preference Rankings. Based on ordinal data, where participants rated interaction techniques from best to worst.

4 Study 1: Results

This section presents the results from Study 1, where interaction techniques (Press, Hover, Head, Eye) are compared for making selections in the Intimate proxemic zone (<0.5m). As all data followed a non-normal distribution, Aligned Rank Transform (ART) repeated-measures ANOVAs [65] were applied (participant random intercept; Kenward-Roger dfs) with Bonferroni correction, reporting partial η^2 effect sizes.

4.1 Selection Time

Selection times were based on individual trials as defined in Section 3.7. Trials were removed where participants made an incorrect selection (n = 8, 0.69%). Outliers were also omitted (n = 16, 1.39%), which were defined as any attempts above three standard deviations from mean selection times (mean ± 3std.).

When analysing the resulting 1,128 trials, a significant main effect was found between technique and selection time ($F_{3,93}$ = 33.669, $p < .001$, $\eta_p^2 = .521$). Press was faster than all techniques ($p < .001$), with Hover having the worst performance ($p < .001$). No significant differences were found between Head and Eye. Figure 3 presents an overview of selection times for each technique.

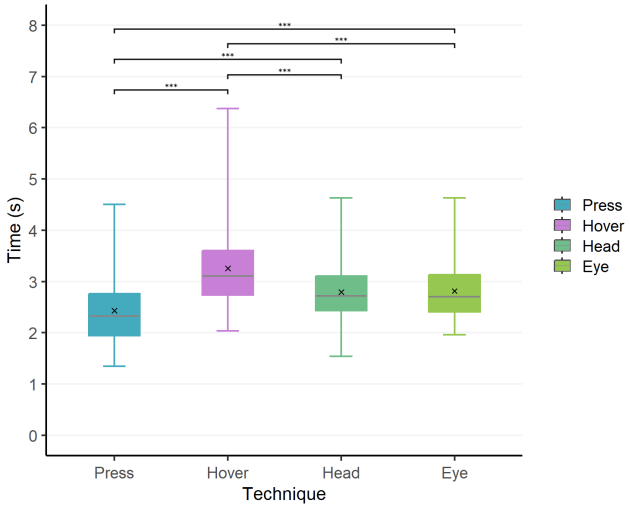


Figure 3: Boxplot showing distribution of selection times with significant differences for Press, Hover, Head and Eye in the intimate proxemic zone. Mean times are indicated by 'X'. Outliers were not included.

4.2 Error

Error rate was calculated based on the number of trials where an incorrect selection was triggered before a correct selection (N = 8). Hover produced the most errors within the intimate proxemic zone (N = 5), however no significant effects were found. Press, Head and Eye produced one error each.

4.3 Task Load

Eye was found to require the least task load (M = 25.57, SD = 16.34), followed by Head (M = 29.56, SD = 18.62), Press (M = 30.73, SD = 18.58) and Hover (M = 30.78, SD = 18.82). No significant main effects were found between techniques, however, ART ANOVA tests revealed differences for the Physical ($F_{3,93} = 5.57, p < .01, \eta_p^2 = .152$), and Effort ($F_{3,93} = 4.76, p < .01, \eta_p^2 = .133$) subscales. Post-hoc pairwise comparisons suggest that Eye is less Physically Demanding than Press ($p < .01$) and Hover ($p < .05$). Eye also required less Effort than all other techniques explored ($p < .05$). A breakdown of NASA-TLX results is provided in Figure 5.

4.4 Preference

After completing all selection trials, participants ranked techniques in order of preference. Eye was ranked first most frequently (43.75%), followed by Head (28.13%), Press (21.88%) and Hover (6.25%).

When considering open subjective responses, Eye was often deemed as easy/intuitive (n = 14) - P1:"It seemed really intuitive", P11:"Eye just felt very easy", comfortable/requiring minimal fatigue (n = 7) - P27:"Eye gaze required little effort", and/or efficient (n = 6) - P15:"It's the most efficient, so I'd choose eye". In contrast, although

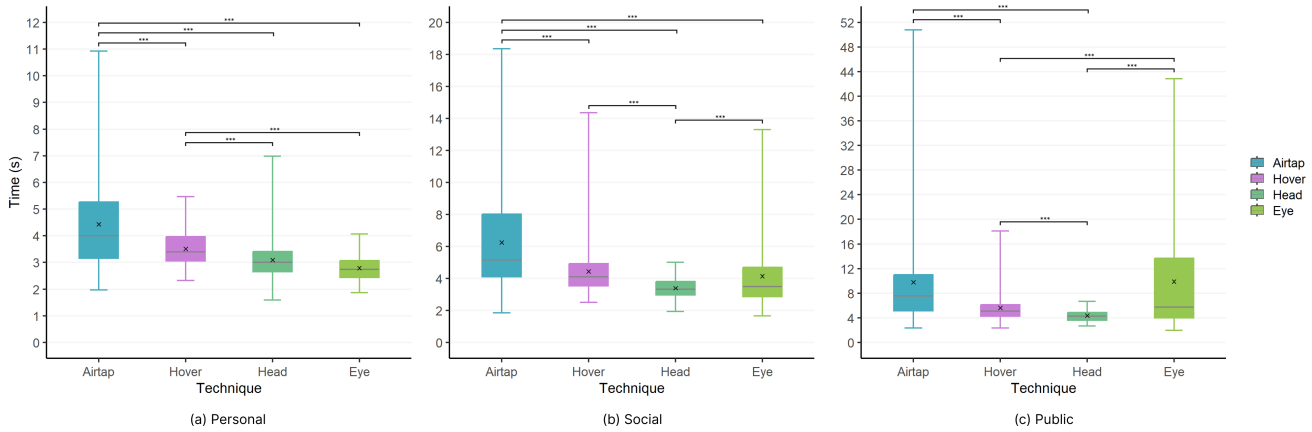


Figure 4: Boxplots showing distribution of selection times with significant differences for Airtap, Hover, Head and Eye in the Personal (a), Social (b) and Public (c) proxemic zones. Mean times are indicated by 'X'. Outliers were not included.

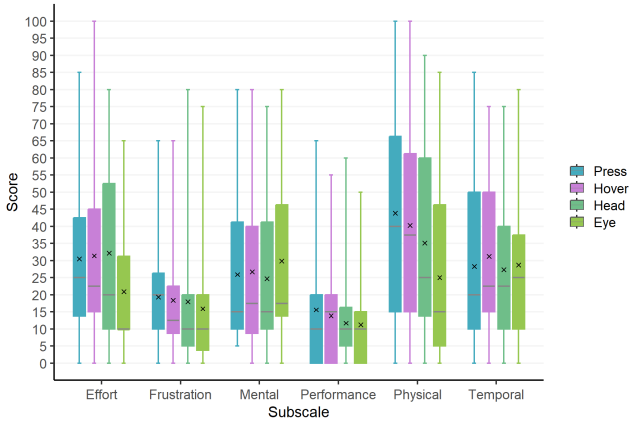


Figure 5: NASA-TLX subscale scores from Study 1 for Press, Hover, Head and Eye within the intimate proxemic zone. Mean scores are indicated by 'X'.

regarded as straightforward to employ (n = 16) - "Really easy, I could just point and select", P13:"You just move your head and it identifies quickly", Head was noted to be more physical/fatiguing (n = 12) - P19:"Head adds physical demand".

Press interaction was viewed as natural/familiar (n = 11) - P2:"Press is more natural, like clicking a real button", however, the fatigue experienced caused an issue for some users (n = 8) - P24:"I kept pushing and it got tiring", whereas Hover was most often regarded as difficult to employ (n = 14) - P15:"The major challenge with hover is that it's hard work using the cursor".

5 Study 2: Results

This section presents the results from Study 2, where each interaction technique (Airtap, Hover, Head, Eye) is compared for performing selections across Personal, Social and Public proxemic zones. As all data again followed a non-normal distribution, Aligned Rank Transform (ART) repeated-measures ANOVAs [65] were applied (participant random intercept; Kenward–Roger dfs) with Bonferroni correction, reporting partial η^2 effect sizes.

5.1 Selection Time

After removing incorrect selections (n = 54, 1.56%) and outliers (n = 56, 1.62%), selection times were evaluated for 3346 trials.

Significant interaction effects were found between technique and distance ($F_{6,341} = 37.927, p < .001, \eta_p^2 = .400$). Post-hoc analysis revealed that, for selections within the Personal proxemic zone, Head and Eye outperformed Hover and Airtap ($p < .001$). Hover also surpassed Airtap ($p < .001$).

At the Social distance, Head outperformed all techniques ($p < .001$). Eye and Hover were also significantly faster than Airtap ($p < .001$). Within the Public proxemic zone, Head again exceeded the other techniques ($p < .001$), with Hover providing faster performance than Airtap and Eye ($p < .001$). Selection times are presented based on proxemic zone in Figure 4.

5.2 Error

Overall, 54 trials resulted in errors (where an incorrect selection was triggered before the correct selection). Statistically significant interaction effects were found between Interaction Technique and Distance ($F_{6,341} = 36.387, p < .001, \eta_p^2 = .369$), with post-hoc analysis revealing that Eye performed significantly worse at the Public distance (n = 33) than every other technique at all distances ($p < .001$). Although Hover produced the most errors in the Personal proxemic zone (n = 5), no other significant differences were found. All other technique/distance comparisons resulted in minimal errors (between 0 and 3).

5.3 Task Load

Head required the least taskload (M = 35.12, SD = 19.93), followed by Hover (M = 38.97, SD = 18.83), Eye (M = 50.00, SD = 23.07) and Airtap (M = 57.05, SD = 16.03). ART ANOVA tests revealed a significant difference between techniques ($F_{3,93} = 16.711, p <$

.001, η_p^2 = .350), with post-hoc pairwise comparisons suggesting that users experienced less overall task load with Head than with Eye ($p < .001$) and Airtap ($p < .001$). Hover was found to require significantly less task load than Airtap ($p < .001$) and Eye ($p < .05$).

Significant differences were also found for the Physical ($F_{3,93} = 9.125, p < .001, \eta_p^2 = .227$), Performance ($F_{3,93} = 11.743, p < .001, \eta_p^2 = .275$), Effort ($F_{3,93} = 13.711, p < .001, \eta_p^2 = .307$) and Frustration ($F_{3,93} = 21.519, p < .001, \eta_p^2 = .410$) subscales. Head and Eye techniques were considered less Physically Demanding than Airtap ($p < .001$), while Head and Hover received lower scores for Frustration than Airtap and Eye ($p < .001$). Head was also considered to require less Effort than Eye and Airtap ($p < .001$), with Hover receiving lower Effort scores than Airtap ($p < .01$). Head scored significantly better than Eye and Airtap in terms of Performance ($p < .001$), with Hover being deemed to have better Performance than Airtap ($p < .01$) and Eye ($p < .05$). A breakdown of NASA-TLX results for each technique is provided in Figure 6.

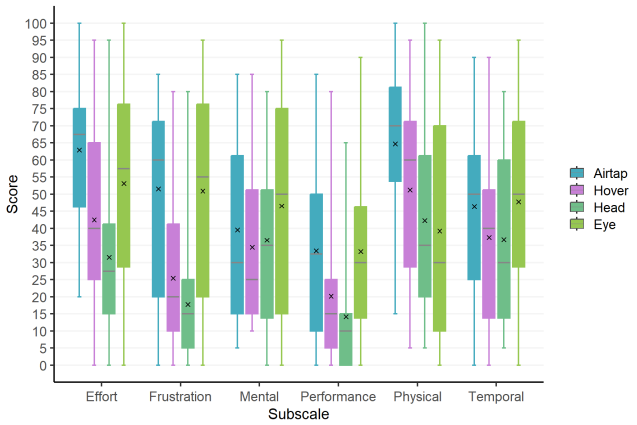


Figure 6: NASA-TLX subscale scores from Study 2 for Airtap, Hover, Head and Eye across Personal, Social and Public proxemic zones. Mean scores are indicated by 'X'.

5.4 Preference

After completing all trials, users ranked the interaction techniques in order of preference. As distances were counterbalanced, participants were asked to provide rankings for objects that were closest to them (Personal) and for objects that were farthest away (Public). Within the Personal zone, Eye was ranked first most often (50%), followed by Head (31.25%), Hover (15.63%) and Airtap (3.13%). When objects were in the Public zone, most participants preferred Head (65.63%), followed by Hover (21.88%), Eye (9.38%) and Airtap (3.13%).

Participants tended to refer to Eye as being quick/easy (n = 12) and accurate (n = 5) when objects were closer - P11: "It was fast and spot-on", P30:"There's very little physical interaction". Conversely, the technique was often considered difficult to employ (n = 9) or the least accurate (n = 8) when objects were placed farther away - P10:"It goes from working to just not working, it falls off a cliff", P21:"It didn't seem like it knew what I was looking at". Head, on the other hand, was reported to be quick/easy (n = 8) and accurate (n = 9) even when selecting objects in the Public zone - P13:"It's consistent, didn't let me down even with smaller targets", P14:"With head (compared to eye) I had a pointer, so I know where I'm aiming".

Hover was deemed more accurate/easier to control than Airtap (n = 8) - P2:"Way quicker than the tap, which seems unnecessary". Many found that Airtap required unnecessary effort for close objects (n = 7) as well as those far away (n = 8), with the technique considered to be awkward/frustrating (n = 6) - P16:"Like why is this (performs gesture) so complicated?", P19:"I had a lot of trouble with that pinch". Unlike other techniques, Head was not ranked last by any participants, highlighting its consistency and reliability - P25:"Head gaze obviously was better", P26:"Head will be winner".

6 Discussion

How we interpret depth and engage with virtual content placed throughout the real world is a key consideration in XR environments, previous research highlighting the influence of distance on interaction suitability [24, 63]. Our results have also demonstrated that distance significantly influences the performance and usability of XR interaction, highlighting the potential of designing techniques that adapt based on proxemic dimensions.

6.1 Preference does not Always Align with Performance

While Press yielded the best performance in Study 1, likely due to the control afforded by direct selection and lack of time-based constraints [33, 42], it was not the most preferred.

Direct freehand input like Press is effective when objects are within reach [47]. However, eye gaze is often more suitable for initial target identification, as our gaze naturally aligns with intent to interact [33, 44]. This meant Eye was preferred, requiring significantly less Physical Demand and Effort.

Head and Eye naturally support search tasks and require less physical strain than freehand inputs [24]. Although prior work suggests that task load could have been reduced if content was positioned relative to nearby surfaces, such as the desk [8], in the low-effort "Observe and Interact" scenario explored, users favoured Eye for its comfort, even at the cost of lower performance.

Head was rated as requiring more effort than Eye when objects were closer, likely due to the need for larger head movements to position the cursor [5], whereas Eye allowed for minimal motion if the content was already in view. Overall, this highlights the importance of interaction context and user experience over raw performance in XR interface design.

6.2 Head Gaze is the Most Reliable Technique

Head consistently provided faster selection times than Airtap and Hover, showing reliability across all distances in line with prior work [48]. High speed is attributed to combining visual search and selection into a single action, unlike Airtap and Hover, which require separate visual and manual inputs.

Unlike Head, the performance of Eye declined with distance due to smaller visual angles [39], lack of a visible cursor [60], and system tracking limitations [24]. Participants often compensated physically (e.g., leaning or squinting) and found Eye less usable in the Public zone, where Hover surpassed its performance. While Eye was preferred in Intimate and Personal zones, and Head in the

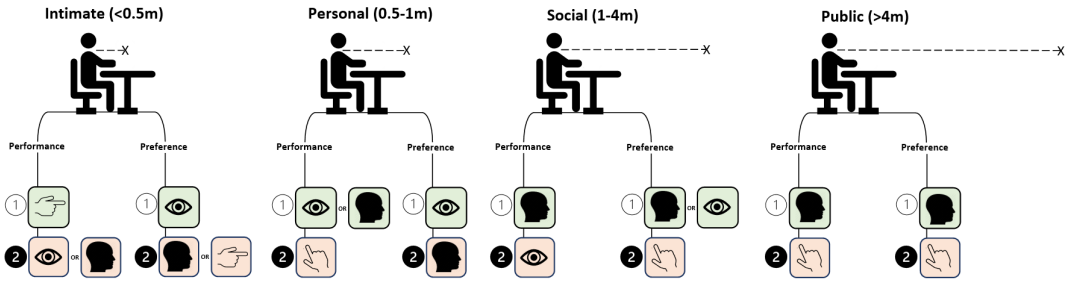


Figure 7: Appropriateness of interaction techniques in different proxemic zones: These are Press, Eye, Head or Hover depending on the distance.

Public zone [59, 63], Eye's higher error rate and lack of feedback reduced user confidence. Improved system feedback could therefore enhance Eye's usability and alignment with user expectations [60].

Airtap, though the baseline technique, suffered from poor flexibility and gesture recognition. Users adapted their gestures for comfort, but input often failed or moved the cursor off target [29, 45]. These issues, tied to system limitations, might be addressed through trajectory-based input [46] and increased interaction zones [66].

6.3 Hover Provided better Usability than Airtap

Hover outperformed Airtap in performance and preference at all distances explored in Study 2, with task load comparable to Head; the highest-ranked technique. This suggests that separating pointing and selection in freehand interaction (e.g. pairing pointing with blink, speech, or a button press) can reduce cognitive load and improve usability [29, 31, 63].

Although Airtap may reduce accidental selections [4], its added complexity and steeper learning curve [45, 48] led to higher frustration. Even among experienced XR users, Hover was generally preferred over Airtap for its simplicity and lower task demand, reinforcing the benefits of simplifying or modularising interaction mechanisms for selection tasks [42].

6.4 Distance could be Referenced for Adapting XR Techniques

Results highlight that distance and angular size are critical in determining the suitability of XR selection techniques [21, 63]. Press worked best for close (Intimate) interactions, reflecting familiar physical behaviours [46], while Airtap was least effective across all distances, requiring more effort and feeling less intuitive [45, 47, 48].

For the task explored, Head or Eye were consistently preferred over freehand methods. Eye was preferred at close range, while Head was more effective at mid-to-far distances, with a switching threshold near the Personal-Social boundary (1.4m), aligning with HoloLens 2's recommendation for a 2° minimum target size [39].

Rather than using one technique across all zones [8], results suggest that dynamically switching inputs based on user-object distance can improve usability and flow. This is especially useful for continuous, multi-purpose XR environments [16], which call for adaptive, context-aware interactions [30]. Figure 7 outlines these recommendations, emphasising the importance of spatial context in XR interaction design.

7 Limitations and Future work

While the studies provide a useful foundation for distance-based adaptive XR interactions, their scope is limited by the technologies, tasks, and user group examined. Focus is also given to performance outcomes, giving less attention to critical experiential factors such as presence, immersion, naturalness, and engagement [1, 11], which limits broader insights into user experience.

Future research should consider defining distance thresholds that are responsive to diverse user needs and usage contexts, especially as traditional proxemic theories may not translate directly into XR environments [20]. There is a need to explore how proxemic adaptation can be combined with approaches like AI-driven interaction prediction [9, 30] or intent-based selection [44]. It is also essential that systems communicate these adaptations clearly to users, particularly during input transitions, to prevent cognitive overload [60]. Designing intuitive and responsive feedback mechanisms will be crucial to ensure user trust and smooth interactions.

While this work has focused on unimodal input methods typical of current commercial XR systems, incorporating multimodal approaches—such as combining gaze, gesture, and voice—could significantly enhance interaction efficiency and usability [24, 31]. Since selection also often precedes more complex actions like object manipulation [64], our findings may help guide the design of a range of distance-sensitive, multimodal interaction techniques.

Furthermore, given that studies were conducted in seated conditions, future work should explore how distance impacts interaction when users are standing or walking [25, 70]. Expanding the investigation of proxemic dimensions, to consider factors such as orientation, movement, identity, and location, could unlock more adaptive and personalised XR experiences [2, 15].

Safe and Responsible Innovation Statement. We underscore the importance of ethically managing the complexities of capturing, interpreting, and adapting XR experiences based on human, environmental and system factors [16]. Recognising the challenges of accurately defining context, e.g. through identity descriptors and spatial/semantic understanding of dynamic spaces [9, 30, 54, 69], this work promotes transparent, secure, and user-centric data practices [67]. Guided by privacy and security [34, 53, 68], we advocate adaptive interactions that balance system responsiveness with user autonomy, contributing responsibly to the advancement of more intelligent XR technologies.

8 Conclusion

This paper has highlighted distance to be a key contextual cue in XR interaction, and demonstrated the value of using proxemic dimensions to inform adaptive interaction design. Findings from two studies show that interaction performance varies by proxemic zone. Head was found to be the most reliable technique overall, with Eye performing well and being most preferred when content is closer to the user. While Press worked best in terms of performance at arm's reach, it was found tiring, and Airtap proved least practical at all distances considered. These results support the need for spatially adaptive, context-aware systems, offering actionable insights into distance-based selection techniques. Findings suggest that harnessing proxemic factors could offer a way to adapt to individual users, their preferences and a range of dynamic environments, providing a potential route towards personalised, AI-driven XR experiences.

References

[1] Huidong Bai, Prasanth Sasikumar, Jing Yang, and Mark Billinghurst. 2020. A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* CHI 2020, April 25–30, 2020, Honolulu, HI, USA (04 2020). doi:10.1145/3313831.3376550

[2] Till Ballendat, Nicolai Marquardt, and Saul Greenberg. 2010. Proxemic interaction: designing for a proximity and orientation-aware environment. *ACM International Conference on Interactive Tabletops and Surfaces - ITS '10* (2010), 121–130. doi:10.1145/1936652.1936676

[3] Ana M. Bernardos, David Gómez, and José R. Casar. 2016. A Comparison of Head Pose and Deictic Pointing Interaction Methods for Smart Environments. *International Journal of Human-Computer Interaction* 32 (01 2016), 325–351. doi:10.1080/10447318.2016.1142054

[4] Shimmila Bhowmick, Pratul Kalita, and Keyur Sorathia. 2020. A Gesture Elicitation Study for Selection of Nail Size Objects in a Dense and Occluded Dense HMD-VR. *IndiaHCI '20: Proceedings of the 11th Indian Conference on Human-Computer Interaction* (11 2020), 12–23. doi:10.1145/3429290.3429292

[5] Jonas Blattgerste, Patrick Renner, and Thies Pfeiffer. 2018. Advantages of eye-gaze over head-gaze-based selection in virtual and augmented reality under varying field of views. *Proceedings of the Workshop on Communication by Gaze Interaction* (06 2018). doi:10.1145/3206343.3206349

[6] Doug A. Bowman and Larry F. Hodges. 1999. Formalizing the Design, Evaluation, and Application of Interaction Techniques for Immersive Virtual Environments. *Journal of Visual Languages & Computing* 10 (02 1999), 37–53. doi:10.1006/jvlc.1998.0111

[7] Ariel Caputo, Riccardo Bartolomioli, and Andrea Giachetti. 2023. Remote and Deviceless Manipulation of Virtual Objects in Mixed Reality. In *Smart Tools and Applications in Graphics - Eurographics Italian Chapter Conference*, Francesco Banterle, Giuseppe Caggianese, Nicola Capece, Ugo Erra, Katia Lupinetti, and Gilda Manfredi (Eds.). The Eurographics Association. doi:10.2312/stag.20231290

[8] Yi Fei Cheng, Christoph Gebhardt, and Christian Holz. 2023. InteractionAdapt: Interaction-driven Workspace Adaptation for Situated Virtual Reality Environments. *ACM Symposium on User Interface Software and Technology (UIST '23)* (10 2023). doi:10.1145/3586183.3606717

[9] Shakiba Davari, Daniel Stover, Alexander Giovannelli, Cory Ilo, and Doug A Bowman. 2024. Towards Intelligent Augmented Reality (iAR): A Taxonomy of Context, an Architecture for iAR, and an Empirical Study. https://arxiv.org/abs/2411.02684

[10] Marcos Daza, Dennis Barrios-Aranibar, José Diaz-Amado, Yudith Cardinale, and João Vilasboas. 2021. An Approach of Social Navigation Based on Proxemics for Crowded Environments of Humans and Robots. *Micromachines* 12 (02 2021), 193. doi:10.3390/mi12020193

[11] Carlos Flavián, Sergio Ibáñez Sánchez, and Carlos Orús. 2018. The impact of virtual, augmented and mixed reality technologies on the customer experience. *Journal of Business Research* 100 (11 2018), 547–560. doi:10.1016/j.jbusres.2018.10.050

[12] Maite Frutos-Pascual, Chris Creed, and Ian Williams. 2019. Head Mounted Display Interaction Evaluation: Manipulating Virtual Objects in Augmented Reality. *Human-Computer Interaction – INTERACT 2019* 11749 (2019), 287–308. doi:10.1007/978-3-030-29390-1_16

[13] Andrea Gallardo, Chris Choy, Jaideep Juneja, Efe Bozkir, Camille Cobb, Lujo Bauer, and Lorrie Cranor. 2023. Speculative Privacy Concerns about AR Glasses Data Collection. *Proceedings on Privacy Enhancing Technologies* 2023 (10 2023), 416–435. doi:10.56553/popets-2023-0117

[14] Saul Greenberg, Sebastian Boring, Jo Vermeulen, and Jakub Dostal. 2014. Dark patterns in proxemic interactions. *Proceedings of the 2014 conference on Designing interactive systems - DIS '14* (2014). doi:10.1145/2598510.2598541

[15] Saul Greenberg, Nicolai Marquardt, Till Ballendat, Rob Diaz-Marino, and Miaosen Wang. 2011. Proxemic interactions. *interactions* 18 (01 2011), 42. doi:10.1145/1897239.1897250

[16] Jens Grubert, Tobias Langlotz, Stefanie Zollmann, and Holger Regenbrecht. 2017. Towards Pervasive Augmented Reality: Context-Awareness in Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 23 (06 2017), 1706–1724. doi:10.1109/tvcg.2016.2543720

[17] Edward T Hall. 1966. *The hidden dimension.* Anchor Books.

[18] Julia Hertel, Sukran Karaosmanoglu, Susanne Schmidt, Julia Braker, Martin Semmann, and Frank Steinicke. 2021. A Taxonomy of Interaction Techniques for Immersive Augmented Reality based on an Iterative Literature Review. *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (10 2021). doi:10.1109/ismar52148.2021.00060

[19] Jinghui Hu, John J Dudley, and Per Ola Kristensson. 2022. An Evaluation of Caret Navigation Methods for Text Editing in Augmented Reality. *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (10 2022). doi:10.1109/ismar-adjunct57072.2022.00132

[20] Ann Huang, Pascal Knierim, Francesco Chiossi, Lewis L Chuang, and Robin Welsch. 2022. Proxemics for Human-Agent Interaction in Augmented Reality. *CHI Conference on Human Factors in Computing Systems* (04 2022), 1–13. doi:10.1145/3491102.3517593

[21] Muhammad Hussain, Jaehyun Park, and Hyun K Kim. 2023. Effects of Interaction Method, Size, and Distance to Object on Augmented Reality Interfaces. *Interacting with computers* 35 (01 2023), 1–11. doi:10.1093/iwc/iwad034

[22] The NASA TLX Tool: Task Load Index. 2020. TLX NASA Ames - Home. https://humansystems.arc.nasa.gov/groups/tlx/

[23] Hyo Jeong Kang, Jung-hye Shin, and Kevin Ponto. 2020. A Comparative Analysis of 3D User Interaction: How to Move Virtual Objects in Mixed Reality. 275–284 pages. doi:10.1109/VR46266.2020.00047

[24] Mikko Kytö, Barrett Ens, Thammathip Piumsomboon, Gun A. Lee, and Mark Billinghurst. 2018. Pinpointing. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (04 2018). doi:10.1145/3173574.3173655

[25] Wallace S. Lages and Doug A. Bowman. 2019. Walking with adaptive augmented reality workspaces. *Proceedings of the 24th International Conference on Intelligent User Interfaces* (03 2019). doi:10.1145/3301275.3302278

[26] Gergana Lilligreen, Nico Henkel, and Alexander Wiebel. 2022. Near and Far Interaction for Augmented Reality Tree Visualization Outdoors. *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* (01 2022). doi:10.5220/0010785700003124

[27] Lars Lischke, Sven Mayer, Katrin Wolf, Niels Henze, Harald Reiterer, and Albrecht Schmidt. 2016. Screen arrangements and interaction areas for large display work places. *KOPS (University of Konstanz)* (06 2016). doi:10.1145/2914920.2915027

[28] Chang Liu, Alexander Plopski, and Jason Orlosky. 2020. OrthoGaze: Gaze-based three-dimensional object manipulation using orthogonal planes. *Computers & Graphics* 89 (06 2020), 1–10. doi:10.1016/j.cag.2020.04.005

[29] Xinyi Liu, Xuanru Meng, Becky Spittle, Wenge Xu, BoYu Gao, and Hai-Ning Liang. 2022. Exploring Text Selection in Augmented Reality Systems. *Proceedings of the 18th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry* (12 2022). doi:10.1145/3574131.3574459

[30] Xingyu Bruce Liu, Jiahao Nick Li, David Kim, Xiang Anthony Chen, and Ruofei Du. 2024. Human I/O: Towards a Unified Approach to Detecting Situational Impairments. *arXiv (Cornell University)* (05 2024). doi:10.1145/3613904.3642065

[31] Feiyu Lu, Leonardo Pavanatto, and Doug A Bowman. 2023. In-the-Wild Experiences with an Interactive Glanceable AR System for Everyday Use. *SUI '23: Proceedings of the 2023 ACM Symposium on Spatial User Interaction* (10 2023). doi:10.1145/3607822.3614515

[32] Mathias N Lystbæk, Ken Pfeuffer, Jens Emil Grønbæk, and Hans Gellersen. 2022. Exploring Gaze for Assisting Freehand Selection-based Text Entry in AR. *Proceedings of the ACM on human-computer interaction* 6 (05 2022), 1–16. doi:10.1145/3530882

[33] Mathias N. Lystbæk, Peter Rosenberg, Ken Pfeuffer, Jens Emil Grønbæk, and Hans Gellersen. 2022. Gaze-Hand Alignment: Combining Eye Gaze and Mid-Air Pointing for Interacting with Menus in Augmented Reality. *Proceedings of the ACM on Human-Computer Interaction* 6 (05 2022), 1–18. doi:10.1145/3530886

[34] Mark Mcgill. 2021. White Paper - The IEEE Global Initiative on Ethics of Extended Reality (XR) Report–Extended Reality (XR) and the Erosion of Anonymity and Privacy. https://ieeexplore.ieee.org/servlet/opac?punumber=9619997

[35] Mark McGill, Aidan Kehoe, Euan Freeman, and Stephen Brewster. 2020. Expanding the Bounds of Seated Virtual Workspaces. *ACM Transactions on Computer-Human Interaction* 27 (05 2020), 1–40. doi:10.1145/3380959

[36] Microsoft. 2021. Comfort - Mixed Reality. https://learn.microsoft.com/en-us/windows/mixed-reality/design/comfort

[37] Microsoft. 2022. HoloLens environment considerations. https://learn.microsoft.com/en-us/hololens/hololens-environment-considerations

[38] Microsoft. 2022. MRTK2-Unity Developer Documentation - MRTK 2. https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2

[39] Microsoft. 2023. Eye-gaze and dwell - Mixed Reality. https://learn.microsoft.com/en-us/windows/mixed-reality/design/gaze-and-dwell-eyes

[40] Microsoft. 2023. Eye-gaze-based interaction - Mixed Reality. https://learn.microsoft.com/en-us/windows/mixed-reality/design/eye-gaze-interaction

[41] Microsoft. 2023. Interactable object - Mixed Reality. https://learn.microsoft.com/en-us/windows/mixed-reality/design/interactable-object

[42] Aunnoy K Mutasim, Anil Ufuk Batmaz, and Wolfgang Stuerzlinger. 2021. Pinch, Click, or Dwell: Comparing Different Selection Techniques for Eye-Gaze-Based Pointing in Virtual Reality. *ACM Symposium on Eye Tracking Research and Applications* (05 2021). doi:10.1145/3448018.3457998

[43] Alexander Ng, Daniel Medeiros, Mark McGill, Julie R Williamson, and Stephen Brewster. 2021. The Passenger Experience of Mixed Reality Virtual Display Layouts in Airplane Environments. *Enlighten: Publications (The University of Glasgow)* (10 2021). doi:10.1109/ismar52148.2021.00042

[44] Ken Pfeuffer, Yasmeen Abdrabou, Augusto Esteves, Radiah Rivu, Yomna Abdelrahman, Stefanie Meitner, Amr Saadi, and Florian Alt. 2021. ARtention: A design space for gaze-adaptive user interfaces in augmented reality. *Computers & Graphics* (01 2021). doi:10.1016/j.cag.2021.01.001

[45] Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, and Hans Gellersen. 2017. Gaze + pinch interaction in virtual reality. *Proceedings of the 5th Symposium on Spatial User Interaction* (10 2017). doi:10.1145/3131277.3132180

[46] Tran Pham, Jo Vermeulen, Anthony Tang, and Lindsay MacDonald Vermeulen. 2018. Scale Impacts Elicited Gestures for Manipulating Holograms. *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18* (2018). doi:10.1145/3196709.3196719

[47] Thammathip Piumsomboon, David Altimira, Hyungon Kim, Adrian Clark, Gun Lee, and Mark Billinghurst. 2014. Grasp-Shell vs gesture-speech: A comparison of direct and indirect natural interaction techniques in augmented reality. *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (09 2014). doi:10.1109/ismar.2014.6948411

[48] Majid Pourmemar and Charalambos Poullis. 2019. Visualizing and Interacting with Hierarchical Menus in Immersive Augmented Reality. *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry* (11 2019), 1–9. doi:10.1145/3359997.3365693

[49] Yuan Yuan Qian and Robert J. Teather. 2017. The Eyes Don't Have It: An Empirical Comparison of Head-Based and Eye-Based Selection in Virtual Reality. *Proceedings of the 5th Symposium on Spatial User Interaction* (10 2017). doi:10.1145/3131277.3132182

[50] Shirin Sadri, Shalva A. Kohen, Carmine Elvezio, Shawn H. Sun, Alon Grinshpoon, Gabrielle J. Loeb, Naomi Basu, and Steven K. Feiner. 2019. Manipulating 3D Anatomic Models in Augmented Reality: Comparing a Hands-Free Approach and a Manual Approach. *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (10 2019), 93–102. doi:10.1109/ismar.2019.00-21

[51] Maximilian Speicher. 2015. What is Usability? A Characterization based on ISO 9241-11 and ISO/IEC 25010. (02 2015). doi:10.48550/arxiv.1502.06792

[52] Becky Spittle, Maite Frutos-Pascual, Chris Creed, and Ian Williams. 2022. A Review of Interaction Techniques for Immersive Environments. *IEEE Transactions on Visualization and Computer Graphics* (2022), 1–1. doi:10.1109/TVCG.2022.3174805

[53] Marvin Strauss, Viktorija Paneva, Florian Alt, and Stefan Schneegass. 2024. Designing and Evaluating Scalable Privacy Awareness and Control User Interfaces for Mixed Reality. *arXiv (Cornell University)* (09 2024). doi:10.48550/arxiv.2409.00739

[54] Jannis Strecker, Khakim Akhunov, Federico Carbone, Kimberly García, Kenan Bektaş, Andres Gomez, Simon Mayer, and Kasim Sinan Yildirim. 2023. MR Object Identification and Interaction. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 7 (09 2023), 1–26. doi:10.1145/3610879

[55] Ying-Chao Tung, Chun-Yen Hsu, Han-Yu Wang, Silvia Chyou, Jhe-Wei Lin, Pei-Jung Wu, Andries Valstar, and Mike Y. Chen. 2015. User-Defined Game Input for Smart Glasses in Public Space. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15* (2015), 3327–3336. doi:10.1145/2702123.2702214

[56] Jesus Ugarte, Nahal Norouzi, Austin Erickson, Gerd Bruder, and Greg Welch. 2022. Distant Hand Interaction Framework in Augmented Reality. *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (03 2022). doi:10.1109/vrw55335.2022.00332

[57] International Telecommunication Union. 2023. Recommendation ITU-R BT.500-15: Methodologies for the subjective assessment of the quality of television images. https://www.itu.int/rec/R-REC-BT.500-15-202305-I/en

[58] Antonio Emmanuele Uva, Michele Fiorentino, Vito Modesto Manghisi, Antonio Boccaccio, Saverio Debernardis, Michele Gattullo, and Giuseppe Monno. 2019. A User-Centered Framework for Designing Midair Gesture Interfaces. *IEEE Transactions on Human-Machine Systems* 49 (10 2019), 421–429. doi:10.1109/thms.2019.2919719

[59] Stephen Uzor and Per Ola Kristensson. 2021. An Exploration of Freehand Crossing Selection in Head-Mounted Augmented Reality. *ACM Transactions on Computer-Human Interaction* 28 (10 2021), 1–27. doi:10.1145/3462546

[60] Roshan Venkatakrishnan, Rohith Venkatakrishnan, Balagopal Raveendranath, Christopher C Pagano, Andrew C Robb, Wen-Chieh Lin, and Sabarish V Babu. 2023. Give Me a Hand: Improving the Effectiveness of Near-field Augmented Reality Interactions By Avatarizing Users' End Effectors. *IEEE Transactions on Visualization and Computer Graphics* 29 (05 2023), 2412–2422. doi:10.1109/tvcg.2023.3247105

[61] Gang Wang, Gang Ren, Xinye Hong, Xun Peng, Wenbin Li, and Eamonn O'Neill. 2022. Freehand Gestural Selection with Haptic Feedback in Wearable Optical See-Through Augmented Reality. *Information* 13 (12 2022), 566. doi:10.3390/info13120566

[62] Tianyi Wang, Xun Qian, Fengming He, Xiyun Hu, Yuanzhi Cao, and Karthik Ramani. 2021. GesturAR: An Authoring System for Creating Freehand Interactive Augmented Reality Applications. *The 34th Annual ACM Symposium on User Interface Software and Technology* (10 2021). doi:10.1145/3472749.3474769

[63] Matt Whitlock, Ethan Harnner, Jed R. Brubaker, Shaun Kane, and Danielle Albers Szafir. 2018. Interacting with Distant Objects in Augmented Reality. *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (03 2018), 42–48. doi:10.1109/vr.2018.8446381

[64] Adam S. Williams, Jason Garcia, and Francisco Ortega. 2020. Understanding Multimodal User Gesture and Speech Behavior for Object Manipulation in Augmented Reality Using Elicitation. *IEEE Transactions on Visualization and Computer Graphics* 26 (12 2020), 3479–3489. doi:10.1109/tvcg.2020.3023566

[65] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. *CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (05 2011), 143–146. doi:10.1145/1978942.1978963

[66] Wenge Xu, Hai-Ning Liang, Yuzheng Chen, Xiang Li, and Kangyou Yu. 2020. Exploring Visual Techniques for Boundary Awareness During Interaction in Augmented Reality Head-Mounted Displays. 204–211 pages. doi:10.1109/VR46266.2020.00039

[67] Xuhai Xu, Mengjie Yu, Tanya R Jonker, Kashyap Todi, Feiyu Lu, Xun Qian, Marcelo Evangelista, Tianyi Wang, Michelle Li, Aran Mun, Te-Yen Wu, Junxiao Shen, Ting Zhang, Narine Kokhlikyan, Fulton Wang, Paul Sorenson, Sophie Kahyun Kim, and Hrvoje Benko. 2023. XAIR: A Framework of Explainable AI in Augmented Reality. *arXiv (Cornell University)* (04 2023). doi:10.1145/3544548.3581500

[68] Muhammad Zahid Iqbal, Xuanhui Xu, Vivek Nallur, Mark Scanlon, and Abraham G Campbell. 2023. Security, Ethics and Privacy Issues in the Remote Extended Reality for Education. *Gaming media and social effects* (01 2023), 355–380. doi:10.1007/978-981-99-4958-8_16

[69] Bowen Zhang, Zhijin Qin, Yiyu Guo, and Geoffrey Ye Li. 2022. Semantic Sensing and Communications for Ultimate Extended Reality. *arXiv (Cornell University)* (01 2022). doi:10.48550/arxiv.2212.08533

[70] Qiushi Zhou, Difeng Yu, Martin N Reinoso, Joshua Newn, Jorge Goncalves, and Eduardo Velloso. 2020. Eyes-free Target Acquisition During Walking in Immersive Mixed Reality. *IEEE Transactions on Visualization and Computer Graphics* 26 (12 2020), 3423–3433. doi:10.1109/tvcg.2020.3023570