Stretch Gaze Targets Out: Experimenting with Target Sizes for Gaze-Enabled Interfaces on Mobile Devices

Omar Namnakani
o.namnakani.1@research.gla.ac.uk
University of Glasgow
Glasgow, United Kingdom

Yasmeen Abdrabou
yasmeen.essam@unibw.de
Lancaster University
Lancaster, United Kingdom
Human-Centered Technologies for Learning,
Technical University of Munich
Munich, Germany

Jonathan Grizou
jonathan.grizou@glasgow.ac.uk
University of Glasgow
Glasgow, United Kingdom
GrizAI
Glasgow, United Kingdom

Mohamed Khamis
mohamed.khamis@glasgow.ac.uk
University of Glasgow
Glasgow, United Kingdom

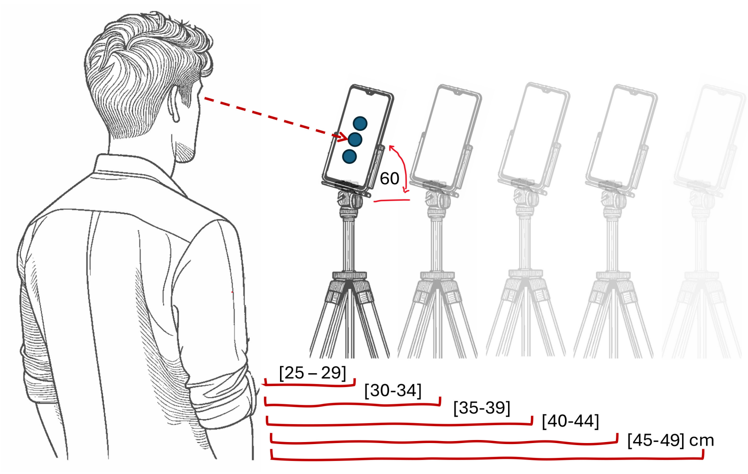

Figure 1: In this work, we investigate the relationship between target size, target region and distance between user and screen. We investigate five target sizes in visual angles, three screen regions, and five face-to-screen distances. The figure shows the five distances used in the study with the mobile phone fixed at 60°. Based on our investigation, we present concrete design guidelines allowing developers to optimise target sizes on gaze-enabled interfaces on mobile devices.

ABSTRACT

Users hold their mobile phones at varying distances depending on their posture, the application being used, and the task's nature. Without considering such variation when designing UI target sizes limits the applicability of gaze selection for everyday interaction with mobile devices. Towards this end, we conducted a user study ($N = 24$) to investigate the implications of different target sizes and viewing across different screen regions. While larger targets generally improve accuracy and decrease precision, accuracy is significantly higher in the horizontal than in the vertical direction. This subsequently led us to find that increasing the tracking area in the vertical direction only, while maintaining the same visual target size, significantly improves accuracy. This suggests that visually smaller targets with larger vertical tracking areas enhance accuracy. Based on our results, we present concrete design guidelines for developers to optimise target sizes on gaze-enabled mobile devices to improve accuracy across varying user-to-screen distances.

CHI '25, April 26-May 1, 2025, Yokohama, Japan

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26-May 1, 2025, Yokohama, Japan*, https://doi.org/10.1145/3706598.3713092.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **Interaction techniques**.

KEYWORDS

Eye Tracking, Gaze-enabled Interfaces, Gaze-based Interaction, Mobile Devices

ACM Reference Format:
Omar Namnakani, Yasmeen Abdrabou, Jonathan Grizou, and Mohamed Khamis. 2025. Stretch Gaze Targets Out: Experimenting with Target Sizes for Gaze-Enabled Interfaces on Mobile Devices. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26-May 1, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3706598.3713092

1 INTRODUCTION

Researchers have explored various gaze input techniques to interact with mobile devices. Dwell time was traditionally one of the early techniques introduced to overcome the Midas touch problem [31]. Other techniques that rely on relative eye movements, such as Pursuits [20, 64] and Gestures [26, 43], were introduced as alternatives to minimise the problem. However, Dwell time interfaces still continue to exist, as they show promise in mobile contexts where they were favoured by users [38, 47]. Also, none of the other proposed techniques provides the same point-and-select functionality as a touchscreen, giving Dwell time a distinct advantage as it replicates this capability [23]. This underlines the importance of improving the usability of Dwell interfaces for target selection, as recent research in gaze-enabled interfaces on mobile devices assessed the usability of Dwell time technique on interfaces that were fundamentally optimised for touch [38, 47], featuring buttons and layouts designed to improve selection accuracy using fingertips [3, 6]. While gaze input can function on these interfaces, its effectiveness is limited because the design is not optimised for gaze input. Dwell input is now supported on Apple devices in their latest operating system, enabling users to interact with their iPhones and iPads using a built-in eye-tracking feature that users can activate (Settings -> Accessibility) [4].

Prior work investigated target size for Dwell interfaces [22, 49]. However, when considering the mobile context, such as how users hold their phones at varying distances and the constraints of limited screen real estate [12, 27, 34], such variation largely limits the applicability of the results of such works, given also the limited screen sizes. Hence, in this work, we investigate the relationship and interplay between target size, viewing distance, and screen region on gaze tracking accuracy for dwell-based input on mobile devices.

Through a controlled lab study ($N = 24$), we collected gaze data as users fixated on targets of different sizes based on their viewing distance. We tested five viewing distance ranges: [25-29], [30-34], [35-39], [40-44], and [45-49] centimetres, accommodating both extremes—from very close, e.g. [25-29] cm, to further away, e.g. [45-49] cm. We also experimented with four different target sizes, represented by visual angles of 2°, 3°, 4°, and 5°, across three screen regions: top, middle, and bottom. This work aims to evaluate several interface designs that facilitate point-and-select tasks using gaze alone, comparing the performance of different target sizes and distance combinations. By analysing these configurations, we aim to uncover the strengths and limitations of each, providing insights into how targets can be designed and optimised for gaze-enabled mobile devices when held at different distances in terms of accuracy and precision.

Our results indicate that targets with a visual angle of 4° serve as an optimal "sweet spot" across all screen regions and face-to-screen distances, minimising tracking errors and achieving an overall tracking accuracy of approximately 70% or higher. Furthermore, we found that horizontal tracking accuracy is higher than the vertical one, suggesting that elongated targets yield significantly better results than uniform ones. To maintain the perception of target size (measured in visual angle) in mobile contexts where the phone's position varies, their on-screen dimensions must be adjusted based on the screen's distance from the user. Thus, in our study, we accounted for this by calculating and adjusting the on-screen target dimensions, using units such as centimetres, based on the phone's position at specific viewing distances (see Figure 1). This ensured that participants perceived the targets at the intended visual angles. We found that such adjustments to targets based on viewing distance resulted in tracking accuracy diminishing for all target sizes as the face-to-screen distance decreased. This paper discusses the interplay between target size, target region, and face-to-screen distance and concludes with guidelines for enabling gaze interactions on mobile devices. Our findings show greater potential for gaze-enabled interfaces and highlight the need for tailored gaze-based interfaces that leverage the strength of gaze input towards more naturalistic and intuitive interactions. This research contributes to improving the integration of gaze interactions on mobile devices by addressing key factors that impact gaze tracking accuracy. It also calls for more tailored interfaces for gaze-based interaction on mobile phones.

Based on our findings, we present concrete design guidelines to allow developers to optimise target sizes on gaze-enabled interfaces on mobile devices for improved accuracy across varying face-to-screen distances. Our guidelines take into consideration both the horizontal and vertical tracking performance and offer specific target size recommendations based on the face-to-screen distance. Developers can refer to the table 3 to find the minimum visual target size needed to maintain a desired level of accuracy, as well as insights on how adjusting the tracking area, especially in the vertical direction, can further improve accuracy.

CONTRIBUTION STATEMENT

This work has two novel contributions. First, we investigate the relationship between target size, target region, and face-to-screen distance. Second, we provide tailored guidelines for designing gaze-enabled interfaces on mobile devices.

2 RELATED WORK

This work builds upon previous research in three areas: a) gaze interactions on mobile devices, b) the challenges associated with gaze interactions in general and on mobile devices in particular, and c) visual angle and target size for gaze-based input. We then conclude this section with a summary of the research gap and how our work contributes to it.

2.1 Gaze-based Interaction Techniques on Mobile Devices

As mobile devices become increasingly powerful, with improved front-facing cameras, they hold significant potential for integrating and adapting eye-tracking technology [17]. This advancement has driven the development of novel input techniques that leverage gaze as a form of interaction on mobile devices [60]. The strength of gaze input is not only that it provides natural interaction [40], but also that it requires less effort compared to other input techniques [36]. Additionally, gaze-based interaction can help overcome various situational impairments, such as operating the phone with one hand, wearing gloves, using the device in the rain, or when touch-based input is generally unavailable [35, 39, 40, 56, 60].

The most common gaze input technique for pointing and selection is Dwell time [31]. Dwell time requires users to briefly fixate on a target to trigger a selection, helping mitigate the "Midas touch" problem–the challenge of distinguishing between gaze used for viewing versus intentional action [17, 23, 46]. However, Dwell time requires target sizes to be optimised to compensate for the inaccuracy of eye tracking [1, 31, 46].

Alternative gaze-based techniques, such as pursuits and gaze gestures, have been explored to overcome these limitations. Pursuits leverages smooth pursuit eye movements, where the user follows a moving object with their gaze [18, 64, 65]. Though successful on small displays [18], this technique was less preferred during walking scenarios, despite its faster selection speed [47]. Furthermore, as the number of selectable targets increases, background correlations often reduce Pursuits' accuracy [64].

The third most commonly used gaze-input technique is gaze gestures [8, 13, 14, 44]. This method requires users to perform eye movements in specific directions [71] or execute a series of gestures with their eyes [14, 55]. Best et al. introduced a rotary-style, gesture-like PIN entry interface using a boundary-crossing approach with a weighted voting scheme for numerals. By using Fitts' Law to model their task, they attributed the rotary style speed advantage to the constant distances between the interface center and any of its numeric pads and also to the elimination of the need for fixation detection [7]. While gaze gestures showed robustness in mobile context [38], it is slower as the number of targets increase [47]. Additionally, using a large or complex set of gestures can be challenging, as they may be difficult to remember and perform [44].

Although various gaze-based interaction techniques for pointing and selection have been introduced in the literature, Dwell time remains the default for gaze-based interaction on most interfaces [19], and specifically mobile devices [38, 47]. This is largely due to its ease of learning and intuitive nature, as well as being the direct gaze counterpart to mouse clicks and touch taps. Therefore, further research is needed to improve pointing and selection in Dwell-based interfaces, particularly in the context of mobile devices, where varying user holding postures and screen layouts introduce new challenges.

2.2 Factors Influencing Gaze Tracking Accuracy

Eye tracking on mobile devices presents several challenges, amplified by the nature of mobile usage. These challenges can be grouped into two categories: 1) gaze estimation issues from both the user and device perspective and 2) interface design. Gaze estimation accuracy is often affected by factors such as face visibility and occlusion [28, 34]. Huang et al. [28] found that full faces were visible in only 30% of collected images. Similarly, Khamis et al. [34] noted that users frequently occlude their faces with their hands or hold their devices in ways that obscure the camera's view, making gaze estimation difficult. This is in addition to the shaky environment nature of mobile devices e.g. during walking or extreme holding postures e.g. slouching (very close) or taking group selfies (very far). Additionally, changing light conditions or facing direct or no light at all further degrades estimation accuracy [22]. Factors such as makeup, face coverings, eye colour, and ethnicity [25] can influence gaze estimation performance.

On the interface side, precise gaze input is still difficult to provide, leading to solutions that often pair gaze input with additional modalities like a mouse [15, 69], pen [51], touch [50, 52, 57], or gestures [53]. However, such approaches limit the hands-free potential of gaze-based interaction and make it dependent on other modalities. Another approach involves estimating the object of interest [45, 59] or detecting user intent [24] to reduce the selection possibilities and improve accuracy. Yet, this requires extensive training data for prediction models, which is not always readily available.

A simpler solution has been to magnify graphical user interfaces [5, 36, 37, 61], but this reduces the number of selectable targets and is often not preferred by users. Another technique involves invisibly expanding targets, the effective area of the target in motor space–allowing gazes near the target to be still registered as selections–without altering its visual size [41]. Therefore, target shape, size, and location become critical design factors in gaze-based interactions.

Several studies have examined how target region [22], size [22, 48, 68], and spacing between targets [42] affect gaze input. which we will expand on in the following section.

Wang et al.[66] studied the effect of visual and motor space sizes (invisible tracking areas) on gaze-based target selection in desktop environments. They found that larger motor spaces led to more accurate target selection and reduced user frustration. They also discovered that the visual space has a feedforward effect on selection speed and modulates the motor space's effectiveness. These findings align with previous work that tested the benefits of target expansion in motor space [41]. Finally, studies have also looked at optimal dwell times for triggering target selection [48], with a recommended dwell time of 600ms for desktop displays.

2.3 Visual Angle and Target Size for Gaze Selection

Prior work showed that targets need to cover at least one degree of visual angle for effective eye tracking. This requirement stems from the limitations of the fovea, the part of the retina responsible for sharp vision. When targets are smaller than this size, users may find it difficult to select them accurately [32, 41].

Several studies have examined target size [22, 48, 68] for dwell-based interfaces and provided recommendations.

Several studies confirmed that larger targets result in improved gaze input accuracy. This is expected as the larger the target, the easier it is to determine that the user is gazing at it. Examples of this include the work of Sakamaki et al. [30], who investigated the effect of feedback, target size, and colour on gaze accuracy in off-screen object selection. While no significant differences were found across audio, visual, and haptic feedback, they identified that larger targets (12 cm) reduced error rates and improved selection speed. Similarly, Stellmach et al. [63] found that larger target sizes consistently improved selection accuracy, even in cases of target overlap. Their study was in the context of exploring five gaze-supported selection techniques using different target sizes and target distances on large public displays.

Other works attempted to identify specific recommended sizes, but limitations in their study designs leave the results inconclusive. For example, the study by Niu et al. [48] determined that a target size of 256 x 256 pixels, the largest in their experiment, significantly improves task accuracy and reduces errors compared to smaller sizes. However, since no larger targets were included in the study, it remains unclear whether this size is truly optimal or if the results simply reflect the principle of "the larger, the better". Penkar et al. [49] conducted two experiments to evaluate the effects of dwell time, button size, and content placement on gaze interaction accuracy, concluding that large button sizes (e.g., 150 pixels/8 cm in diameter) and short dwell times are most effective, as such parameters minimise the Midas Touch problem and help users maintain their gaze more easily on the target. Similar to the previous study, the recommended diameter was the largest in the experiment, making it unclear if this is an optimal size or if more effective sizes could have been found had even larger targets been considered.

In a more structured approach, Feit et al. [22] provided a formal method to determining target sizes for gaze-enabled applications based on tracking accuracy and precision. They recommend that, for robust interaction where at least 95% of gaze points fall within the target area, the target size should be at least 3.28 x 3.78 cm without filtering and 1.9 x 2.35 cm with optimal filtering, accounting for variations in user tracking quality and screen regions.

Given that the mobile setting is different from other setups where the recommendation for target sizes was based on evaluations on desktops at fixed distances between the users and the screens, this reduces the applicability of transferring findings [33]. The user holds their mobile devices at varying distances when interacting with mobile devices [12, 27, 34], and such a factor directly impacts how the user perceives the targets in visual angles, affecting the accuracy of tracking and the overall usability of the interaction.

2.4 Summary of Research Gap and Our Contribution

As seen from the literature, several factors influence gaze tracking accuracy from the gaze estimation side, the user side, and also from the user interface perspective. Surprisingly, very limited to no work addresses factors influencing gaze tracking on mobile devices and all of its affecting factors. In this work, we investigate factors affecting gaze-based selection on mobile devices namely: target size, target screen region, and user-to-screen distance. Toward this end, we aim to answer the following research question: **RQ:** What is the interplay between target size, user-to-screen distance, and target region on the tracking accuracy on mobile devices?

3 STUDY DESIGN AND DATA COLLECTION

To study the implications of target sizes, screen regions, and viewing distances on the tracking accuracy and precision of eye gaze interactions, we collect eye tracking data in a controlled setup to explore how these variations impact user interaction.

3.1 Design

To answer our research question, we created a within-subject design with three independent variables: 1) **Target Size**: We accounted for varying screen sizes by using consistent visual angle measurements for the target sizes, specifically 2°, 3°, 4°, and 5° in both width and height. Such measurements allowed for a maximum of three targets to fit vertically on the screen when using 5°, 2) **Viewing distance**: Given the unstable nature of using a mobile phone, we chose distance ranges rather than fixed distances. We considered five different ranges between the smartphone and the viewer. These distances were chosen based on the typical comfort range for mobile use [27], with the extremes representing very close distances to simulate slouching and the furthest distances to simulate taking selfies. Our choice distances are 25-29cm, 30-34cm, 35-39cm, 40-44cm, and 45-49cm. 3) **Target Region:** We also evaluated the tracking accuracy at three different regions on the phone's screen: Top (nearest to the tracking camera), Middle, and Bottom (farthest from the tracking camera) regions. Each of the three screen regions contains a single centered target, with all three targets always visible. The target to be gazed at is highlighted in a different colour.

3.2 Study Software Implementation

We developed an iOS application and ran it on an iPhone X with a 5.8-inch display. The phone features a 7 MP front camera with an f/2.2 aperture and a 32mm standard focal length. The application utilised the ARKit APIs in iOS to determine the distance between the camera and the participant's face. The size of the targets in centimetres for each level of *Target Size* in visual angle was calculated considering the median viewing distances within the specified ranges of distances outlined in Section 3.1, using the following derived formula [16, 21]:

$$TargetSize = 2 \cdot Distance \cdot \tan\left(\frac{VisualAngle}{2}\right) \tag{1}$$

To enable tracking on the phone using the front-facing camera, we used Eyedid, the commercially available tracking library[1], at 30 frames per second. The Eyedid library provides real-time gaze data recorded as x and y coordinates on the screen, with a reported average accuracy of 1.6°.

3.3 Setup and Apparatus

Our controlled experiment was conducted in a quiet room on the university campus without windows. Participants were seated on

---

[1] Eyedid: https://sdk.eyedid.ai

a chair at a marked spot on the floor, maintaining a straight posture with their backs rested. To simulate natural interactions with mobile devices and to collect accurate gaze data without the noise associated with holding the phone, the mobile device used in the experiment was mounted on a stand with a fixed angle at 60° as done in prior work [27]. The angle was measured using a protractor and Core Motion, an Apple framework with the CMMotionManager class provided to obtain the device motion data, including pitch, roll, and yaw. To control the position of the participants relative to the device, using a visual guide on the display, the stand position and its height were adjusted to ensure the participant's face was visible and centred on the screen while achieving the required face-to-screen distance.

For each level of *Target Size* in visual angle to be maintained and perceived at the intended level at specific viewing distances, we adjusted the phone's position and recomputed the effective on-screen target size, resulting in targets being larger at further distances and smaller at closer distances.

3.4 Participants and Recruitment

Through mailing lists, social media, and word of mouth, we recruited Twenty-four participants (11 female, 13 male) aged 22 to 44 ($M = 29.71, SD = 5.68$). We collected participants' heights to adjust the phone stand accordingly, our participants' heights ranged from $153cm$ to $190cm$, with an average of $167.35cm$ ($SD = 9.79cm$). Out of the 24 participants, 18 reported using vision correction. Among them, two had undergone vision correction surgery, and another two reported using contact lenses. Eight participants wore glasses during the experiment. Eight participants reported suffering from nearsightedness, two were farsighted, one reported astigmatism, and another reported myelitis. Most of the participants reported little experience with eye tracking ($Med = 1.5, Q1 = 0, Q3 = 4$) when responding to a question to rate their experience with eye tracking: "0: no experience—5: very experienced". Those with experience mainly reported using eye tracking in their research with wearable or remote eye trackers, while others had participated in different eye-tracking experiments. Each participant was compensated with a £10 gift voucher for participating. The study received approval from our institute's ethics committee (Approval Number: 300230133).

3.5 Data Collection

Gaze data, recorded as (x,y) coordinates, were captured using Eyedid's API (see Section 3.2). Following practice from prior work, the gaze data were collected for each target during a 1-second window (30 frames) [22], beginning when a participant's gaze entered the target area. A random number was displayed in the center of the target toward the end of the $1s$ interval. Participants were required to enter the number into a prompted text field to confirm their continued focus on the target. This approach ensured participant's active engagement and maintained their attention throughout the collection process. For each condition, we collected gaze data from three different regions of the screen—top, middle, and bottom—when the phone was held in portrait orientation. In each trial, the targets were centred within the specified region.

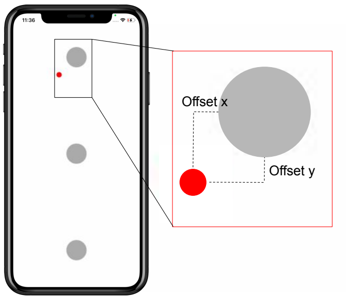

Figure 2: The offset error is computed between the mean of the estimated gaze points and the target area boundary separately for the x and y coordinates. The plot illustrates an example of how the offset error is calculated. This is opposed to *tracking accuracy*, which refers to the percentage of gaze points within the target area, and *precision*, which quantifies the variability of collected gaze data during fixation (See Section 3.7 for measures). The data is from a participant with a *Target Size* of 2° at a *Viewing Distance* of [25-29] cm.

3.6 Task and Procedure

The study started with a consent form and a demographic questionnaire, followed by the tasks. The participant's task was to look at circular targets placed on the screen for 1 second each (see Section 3.5). Participants underwent five blocks, each corresponding to a different viewing distance.

Before each block, we calibrated the phone's front-facing camera with a 5-point calibration method provided by the eye tracking library. **In each block,** participant had to complete one trial for each combination of target size and regions on the screen. Blocks and trials were counterbalanced using Latin square. **Before each trial,** an instruction screen was displayed, showing the size of the target participants had to fixate on, the target region, and the required screen distance, allowing participants to adjust the camera position accordingly, assisted by the experimenter. **In each trial,** three targets appeared at the three different regions on the screen, with the target to be gazed at highlighted with a different background colour. Once the participant gazed inside the highlighted target and the data collection began, the number appeared after $1s$ (see Section 3.5). Upon the time window completion, if the participant entered the correct number in the prompted text field after, the trial ended, and the participant waited for the instruction screen for the subsequent trial to appear. The trial was restarted if the participant did not enter the correct number. We also ended the trial if the participant failed to gaze at the highlighted target after $20s$ to ensure that the experiment ended at a reasonable duration [47]. **After each trial,** in the app, we asked the participant to indicate on a 5-point Likert scale how much they agree with the statement "I found fixating on the target is tiring to the eyes". Each participant

Table 1: Average accuracy for levels of Viewing Distance, Target Size, and Target Region. Overall is the mean accuracy based on the *Target Size* at each *Viewing Distance*. The higher the accuracy, the better. Targets of sizes 4° and 5° achieve an overall tracking accuracy of about 70% or higher as highlighted in dark turquoise. Standard deviations are in parentheses.

Target size [°]	**Region**	**Viewing Distance [cm]**				
		[25-29]	**[30-34]**	**[35-39]**	**[40-44]**	**[45-49]**
2	**Top**	37.4% (40.2)	44.6% (40.5)	51.1% (37.6)	69.6% (33.8)	71.8% (36.7)
3		53.1% (37.9)	66.9% (39.4)	62.5% (37.0)	66.5% (37.8)	78.5% (25.4)
4		70.3% (33.4)	70.1% (33.8)	69.0% (38.4)	79.6% (31.7)	87.8% (25.1)
5		86.4% (25.3)	88.3% (23.4)	82.1% (28.7)	86.5% (25.4)	91.2% (22.4)
2	**Middle**	32.8% (26.4)	44.6% (35.6)	53.8% (39.5)	65.6% (38.1)	59.7% (33.6)
3		52.8% (40.0)	61.9% (38.6)	64.4% (34.9)	86.2% (24.4)	77.1% (30.0)
4		71.1% (32.2)	75.7% (31.3)	83.3% (29.3)	96.7% (6.59)	91.1% (21.8)
5		72.1% (33.7)	89.7% (22.4)	93.2% (18.1)	91.1% (24.3)	95.4% (11.2)
2	**Bottom**	33.5% (33.0)	58.5% (39.6)	55.1% (40.8)	57.9% (38.2)	54.3% (37.0)
3		49.0% (45.1)	65.1% (33.9)	75.0% (30.3)	77.9% (30.3)	70% (30.5)
4		70.6% (38.6)	67.4% (33.4)	83.1% (30.2)	77.4% (35.0)	74.0% (36.6)
5		75.1% (33.0)	77.4% (35.3)	84.9% (25.6)	91.0% (17.0)	85.0% (25.0)
2	**Overall**	34.5% (33.3)	49.2% (38.7)	53.3% (38.8)	64.4% (36.6)	61.9% (36)
3		51.6% (40.6)	64.7% (36.9)	67.3% (34.1)	76.9% (31.9)	75.2% (28.5)
4		70.6% (34.3)	71.1% (32.6)	78.5% (33.1)	84.5% (28.5)	84.3% (29.1)
5		77.9% (31.1)	85.1% (27.8)	86.7% (24.6)	89.5% (22.3)	90.6% (20.6)

completed 4 *Target Sizes* x 5 *Viewing Distances* x 3 *Target Regions* = 60 trials.

3.7 Measures

To analyse our collected data, we set out several dependent variables as measures. The first dependent variable, *tracking accuracy*, is calculated as the percentage of gaze points within the target area when a participant fixates on the highlighted target. This is determined by dividing the number of gaze points inside the region by the total number collected over the 1*s* period. The second dependent variable, *precision*, quantifies the variability of recorded gaze points during fixation by calculating their standard deviation within the 1*s* period as done in prior work [22]. We also computed the *offset error* between the mean of the estimated gaze points and the boundary of the target area separately for the x and y coordinates. The error value is calculated by subtracting the target's radius from the mean of the estimated gaze points along the corresponding axis (see Figure 2).

3.8 Limitation

For this experiment, to enable eye tracking on mobile phones, we used a commercially available eye-tracking library that uses the front-facing cameras of the mobile device; hence, all reported accuracy and precision are based on the accuracy of the library, its calibration method, and the mobile device used for the experiment, including the recommendation for target sizes in Table 3. However, we expect that the relative differences are likely to stay the same when using different phones/tracking software, as prior work suggested that tracking is affected by individuals more than by eye-tracking equipment [22].

Given the limited horizontal screen real estate available on the device, we positioned the targets at the centre of the screen. This allowed us to examine tracking accuracy across three regions, primarily varying in the vertical direction, as larger targets were challenging to fit horizontally in a single row.

4 RESULTS

Our collected dataset consisted of 1440 fixations collected from 24 participants. Fixations refer to where the participants are gazing and looking at a target, and data are collected. We used no algorithm to extract fixations but rather guided participants to gaze at the highlighted target when starting the task for the gaze raw data collection to start. For statistical tests, we used two-way repeated measures ANOVA tests. For Likert scale responses, we used Aligned Rank Transform (ART), a nonparametric equivalent to the factorial ANOVA [67] to transform the data and allow for the analysis of multiple factors. We used Greenhouse-Geisser correction in cases where Mauchly's test indicated a violation of sphericity. To account for multiple comparisons, we corrected P-values for posthoc tests using Bonferroni correction. In scenarios where the participants failed to gaze at the target within 20*s* as done in prior work [47], and a timeout occurred, we recorded the accuracy as 0%.

4.1 Tracking Accuracy

Table 1 shows an overview of the average accuracy results for all screen regions, target size in visual angles, and viewing distance in cm. As expected, the tracking accuracy increases with the target size. Interestingly, it reduces with the closer distance, which is unexpected. This might be because the target sizes are determined by the visual angle. When maintaining the same visual angle but adjusting the phone's position based on viewing distance, the computed on-screen targets are larger at further distances and smaller at closer distances (see Section 4.4). As highlighted in dark turquoise, the table also shows that target sizes of 4° and 5° maintain an overall

Table 2: Significant differences in accuracy across various target sizes in each level of Viewing Distance. Significant pairs (*) are shown for horizontal-vertical directions, horizontal direction, and vertical direction, with mean accuracy and standard deviations reported in brackets. P-values are colour-coded to indicate significance levels, with the lightest shade representing $p < 0.05$ and progressively darker shades for smaller values, specifically $p < 0.01$, $p < 0.005$, $p < 0.001$, $p < 0.0005$, and the darkest shade for $p < 0.0001$.

	Horizontal-Vertical Direction		Horizontal Direction		Vertical Direction	
Viewing Distance	**Target Size [°]**		**Target Size [°]**		**Target Size [°]**	
[25-29]	2 (34.5%, 33.3)	3 (51.6%, 40.6) *	2 (63.9%, 40.4)	3 (83.6%, 30.8) *	2 (46.0%, 35.7)	3 (56.2%, 39.6)
	2 (34.5%, 33.3)	4 (70.6%, 34.3) *	2 (63.9%, 40.4)	4 (95.2%, 17.1) *	2 (46.0%, 35.7)	4 (73.9%, 34.3) *
	2 (34.5%, 33.3)	5 (77.9%, 31.1) *	2 (63.9%, 40.4)	5 (96.2%, 16.7) *	2 (46.0%, 35.7)	5 (79.8%, 30.1) *
	3 (51.6%, 40.6)	4 (70.6%, 34.3) *	3 (83.6%, 30.8)	4 (95.2%, 17.1)	3 (56.2%, 39.6)	4 (73.9%, 34.3)
	3 (51.6%, 40.6)	5 (77.9%, 31.1) *	3 (83.6%, 30.8)	5 (96.2%, 16.7)	3 (56.2%, 39.6)	5 (79.8%, 30.1) *
	4 (70.6%, 34.3)	5 (77.9%, 31.1)	4 (95.2%, 17.1)	5 (96.2%, 16.7)	4 (73.9%, 34.3)	5 (79.8%, 30.1)
[30-34]	2 (49.2%, 38.7)	3 (64.7%, 36.9) *	2 (83.6%, 31.7)	3 (94.8%, 18.3) *	2 (50.9%, 39.0)	3 (65.7%, 36.2) *
	2 (49.2%, 38.7)	4 (71.1%, 32.6) *	2 (83.6%, 31.7)	4 (99.6%, 3.54) *	2 (50.9%, 39.0)	4 (71.5%, 32.0) *
	2 (49.2%, 38.7)	5 (85.1%, 27.8) *	2 (83.6%, 31.7)	5 (99.7%, 2.36) *	2 (50.9%, 39.0)	5 (85.1%, 27.8) *
	3 (64.7%, 36.9)	4 (71.1%, 32.6)	3 (94.8%, 18.3)	4 (99.6%, 3.54)	3 (65.7%, 36.2)	4 (71.5%, 32.0)
	3 (64.7%, 36.9)	5 (85.1%, 27.8) *	3 (94.8%, 18.3)	5 (99.7%, 2.36)	3 (65.7%, 36.2)	5 (85.1%, 27.8) *
	4 (71.1%, 32.6)	5 (85.1%, 27.8) *	4 (99.6%, 3.54)	5 (99.7%, 2.36)	4 (71.5%, 32.0)	5 (85.1%, 27.8) *
[35-39]	2 (53.3%, 38.8)	3 (67.3%, 34.1)	2 (82.0%, 33.2)	3 (94.2%, 20.0)	2 (63.2%, 37.5)	3 (68.2%, 34.3)
	2 (53.3%, 38.8)	4 (78.5%, 33.1) *	2 (82.0%, 33.2)	4 (96.6%, 15.1) *	2 (63.2%, 37.5)	4 (79.8%, 31.3)
	2 (53.3%, 38.8)	5 (86.7%, 24.6) *	2 (82.0%, 33.2)	5 (97.7%, 13.2) *	2 (63.2%, 37.5)	5 (86.9%, 24.5) *
	3 (67.3%, 34.1)	4 (78.5%, 33.1)	3 (94.2%, 20.0)	4 (96.6%, 15.1)	3 (68.2%, 34.3)	4 (79.8%, 31.3)
	3 (67.3%, 34.1)	5 (86.7%, 24.6) *	3 (94.2%, 20.0)	5 (97.7%, 13.2)	3 (68.2%, 34.3)	5 (86.9%, 24.5) *
	4 (78.5%, 33.1)	5 (86.7%, 24.6)	4 (96.6%, 15.1)	5 (97.7%, 13.2)	4 (79.8%, 31.3)	5 (86.9%, 24.5)
[40-44]	2 (64.4%, 36.6)	3 (76.9%, 31.9)	2 (91.6%, 22.0)	3 (95.8%, 17.4)	2 (68.3%, 35.3)	3 (79.1%, 30.8)
	2 (64.4%, 36.6)	4 (84.5%, 28.5) *	2 (91.6%, 22.0)	4 (97.5%, 13.7)	2 (68.3%, 35.3)	4 (84.8%, 28.0) *
	2 (64.4%, 36.6)	5 (89.5%, 22.3) *	2 (91.6%, 22.0)	5 (99.7%, 1.45)	2 (68.3%, 35.3)	5 (89.8%, 22.2) *
	3 (76.9%, 31.9)	4 (84.5%, 28.5)	3 (95.8%, 17.4)	4 (97.5%, 13.7)	3 (79.1%, 30.8)	4 (84.8%, 28.0)
	3 (76.9%, 31.9)	5 (89.5%, 22.3) *	3 (95.8%, 17.4)	5 (99.7%, 1.45)	3 (79.1%, 30.8)	5 (89.8%, 22.2)
	4 (84.5%, 28.5)	5 (89.5%, 22.3)	4 (97.5%, 13.7)	5 (99.7%, 1.45)	4 (84.8%, 28.0)	5 (89.8%, 22.2)
[45-49]	2 (61.9%, 36.0)	3 (75.2%, 28.5)	2 (88.7%, 27.2)	3 (95.4%, 15.5)	2 (66.1%, 35.7)	3 (76.7%, 28.5)
	2 (61.9%, 36.0)	4 (84.3%, 29.1) *	2 (88.7%, 27.2)	4 (96.4%, 17.1)	2 (66.1%, 35.7)	4 (86.5%, 27.0) *
	2 (61.9%, 36.0)	5 (90.6%, 20.6) *	2 (88.7%, 27.2)	5 (100.0%, 0.0) *	2 (66.1%, 35.7)	5 (90.6%, 20.6) *
	3 (75.2%, 28.5)	4 (84.3%, 29.1)	3 (95.4%, 15.5)	4 (96.4%, 17.1)	3 (76.7%, 28.5)	4 (86.5%, 27.0)
	3 (75.2%, 28.5)	5 (90.6%, 20.6) *	3 (95.4%, 15.5)	5 (100.0%, 0.0)	3 (76.7%, 28.5)	5 (90.6%, 20.6) *
	4 (84.3%, 29.1)	5 (90.6%, 20.6)	4 (96.4%, 17.1)	5 (100.0%, 0.0)	4 (86.5%, 27.0)	5 (90.6%, 20.6)

tracking accuracy of at least 70%. The same accuracy is maintained across the vast majority of viewing distances, except for a drop to 69% at [35-39] cm and 67.4% at [30-34] cm for targets of size 4°.

4.1.1 Accuracy per Viewing Distance. For the accuracy based on viewing distance, Table 2 shows the statistical test results. At the *Viewing Distance* of $[25-29]$ cm, there was a significant main effect of *Target Size* on the accuracy, $F_{2.54,\ 58.49}=27.37$, $p < .001$. Overall, targets of size 2° ($M = 34.5\%, SD = 33.3$) were significantly less accurate compared to all other sizes of 3°($M = 51.6\%, SD = 40.6, p < .05$), 4° ($M = 70.6\%, SD = 34.3, p < .0001$), and 5°($M = 77.9\%, SD = 31.1, p < .0001$). Similarly, targets of size 3° were significantly less accurate compared to targets of sizes 4° ($p < .05$) and 5°($p < .005$).

At the *Viewing Distance* of $[30-34]$, there was a significant main effect of *Target Size* on the accuracy, $F_{2.69,\ 61.76}=22.72$, $p < .001$. Overall, targets of size 5° ($M = 85.1\%, SD = 27.8$) were significantly more accurate compared to all other sizes of 2°($M = 49.2\%, SD = 38.7, p < .0001$), 3° ($M = 64.7\%, SD = 36.9, p < .0.001$), and 4°($M = 71.1\%, SD = 32.6, p < .01$). Targets of size 2° were also significantly less accurate than targets of sizes 3° and 4°($p < .0.01$).

At the *Viewing Distances* of $[35-39]$, $[40-44]$, *and*$[45-49]$, there was a significant main effect of *Target Size* on the tracking accuracy, ($F_{2.36,\ 54.20}=13.76$, $p < .001$, $F_{2.54,\ 58.33}=13.66$, $p < .001$, $F_{2.21,\ 50.81}=17.45$, $p < .001$, respectively). Post hoc analysis revealed differences between multiple pairs, shown in Table 2.

While we found no significant main effect of *Target Region* nor an interaction effect between *Target Size* and *Target Region* at almost all levels of *Viewing Distances* ($p > .05$), there was a significant main effect *Target Region* on the tracking accuracy at the *Viewing Distance* of $[45-49]$, $F_{1.97,\ 45.32}=3.86$, $p < .05$. However, post hoc analysis did not reveal significant differences between pairs.

Overall, we found that across all distances, targets of size 4° and 5° are significantly more accurate compared to targets of size 2° (see Table 2). Regardless of the regions, both sizes consistently maintain tracking accuracy above 70% across all *Viewing Distances*. When the distance exceeds 40*cm*, targets of size 3° also achieve tracking accuracy above 70%. This suggests that reducing the visual angle at greater distances could help maintain tracking accuracy while conserving screen space, making room for additional targets.

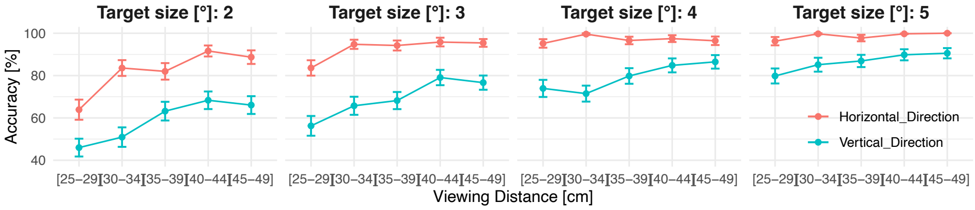

Figure 3: Accuracy based on target fixation in the different viewing distances regardless of the target regions. A paired sample t-test revealed significant differences between the horizontal and vertical directions on the accuracy of the collected gaze data. This demonstrates that accuracy is significantly higher along the horizontal (x-axis) than the vertical (y-axis). Error bars represent standard error.

Observation 1: Targets of sizes 4° and 5° maintain tracking accuracy above 70% across all viewing distances in the horizontal-vertical direction.

4.1.2 Accuracy for Horizontal and Vertical Directions. Figure 3 illustrates the variation in accuracy between the horizontal and vertical direction based on the collected gaze data. A paired sample t-test showed that, on average, accuracy is worse in the vertical direction compared to the horizontal direction t(1439)= 21.626, $p < .0001$.

In the horizontal direction, the *Target Size* had a significant main effect on accuracy at most levels of *Viewing Distance*. Specifically, at *Viewing Distance* of $[25 - 29]$ cm, the effect was $F_{1.75,\ 40.15}$=19.50, $p < .001$; at *Viewing Distance* of [30-34] cm, it was $F_{1.41,\ 32.40}$=11.69, $p < .001$; at *Viewing Distance* of [35-39]cm, it was $F_{1.41,\ 32.41}$= 6.84, $p < .01$; and at *Viewing Distance* of [45-49]cm, it was $F_{2.25,\ 51.75}$= 5.85, $p < .01$. While *Target Region* had a significant main effect on accuracy at *Viewing Distance* of [25-29]cm and [40-44] cm, with $F_{1.97,\ 45.20}$=3.79, $p < .05$, and $F_{1.83,\ 42.10}$=3.75, $p < .05$ respectively, post hoc analysis did not reveal significant differences between pairs, suggesting that placing targets at the top, middle, or bottom regions did not significantly affect the accuracy in the horizontal direction. There was no significant interaction between *Target Size* and *Target Region* at each level of *Viewing Distance*. As the *Viewing Distances* ranges between 25 and 34 cm, pairwise comparison between all levels of *Target Sizes* revealed that targets of size 2° were significantly less accurate compared to targets of all other sizes 3°, 4°, and 5°. Mean values are shown in Table 2. As the *Viewing Distance* increase, ranging between 35 and 49 cm, the notable significant differences were observed between targets of size 2° (M = 82.0%, SD = 33.2) and both 4° (M = 96.6%, SD = 15.1, $p < .05$) and 5° (M = 97.7%, SD = 13.2, $p < .05$) at the *Viewing Distance* of [35-39] cm, and only between targets of size 2° (M = 88.7%, SD = 27.2) and 5° (M = 100%, SD = 0, $p < .01$) at the *Viewing Distance* of [45-49] cm. Overall, while targets of size 4° and 5° achieve a tracking accuracy above 95% in the horizontal direction across all distances, with significantly improved accuracy compared to targets of size 2° at closer viewing distances, targets of size 3° maintain comparable tracking quality as the distance increases ($> 30cm$), with significantly better tracking quality (> 80%) than targets of size 2° at closer distances.

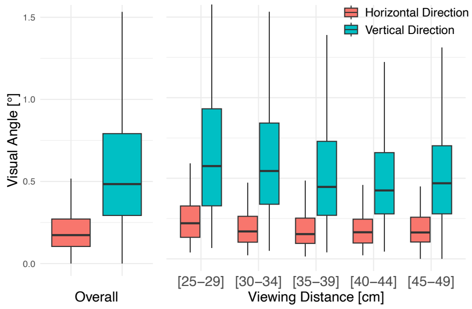

Figure 4: Left: Precision (Standard deviation) of overall target fixation. Right: Precision of target fixation at each viewing distance level. Lower values indicate better precision. Significant differences in precision were found between the horizontal and vertical directions. Horizontal precision consistently outperforms vertical precision across all viewing distances.

Observation 2: As the distance between the viewers and the phone increase, targets with smaller visual angles such as 3° could help maintain good tracking accuracy while conserving screen space.

In the vertical direction, the statistical analysis showed a significant main effect of *Target Size* at all levels of *Viewing Distance*. At *Viewing Distance* of $[25 - 29]$ cm, the effect was $F_{2.65,\ 60.89}$= 16.18, $p < .001$. At *Viewing Distance* of $[30 - 34]$ cm, $F_{2.72,\ 62.51}$= 20.99, $p < .001$. At *Viewing Distance* of $[35 - 39]$ cm, $F_{2.35,\ 54.07}$= 8.68, $p < .001$. At *Viewing Distance* of $[40 - 44]$ cm, $F_{2.64,\ 60.72}$= 11.46, $p < .001$. At *Viewing Distance* of $[45 - 49]$ cm, $F_{2.25,\ 51.75}$= 5.85, $p < .01$. There was no significant main effect of *Target Region*, nor was there an interaction between *Target Size* and *Target Region*. Post hoc analysis revealed significant differences between multiple pairs (see Table 2). Overall, targets of sizes 2° and 3° resulted in significantly less accurate collected gaze data compared to targets of size 5° at most levels of *Viewing distance*. Mean values are shown in Table 2. The results also show that targets of size 4° and 5° achieve a minimum vertical tracking accuracy above 70% at all distances.

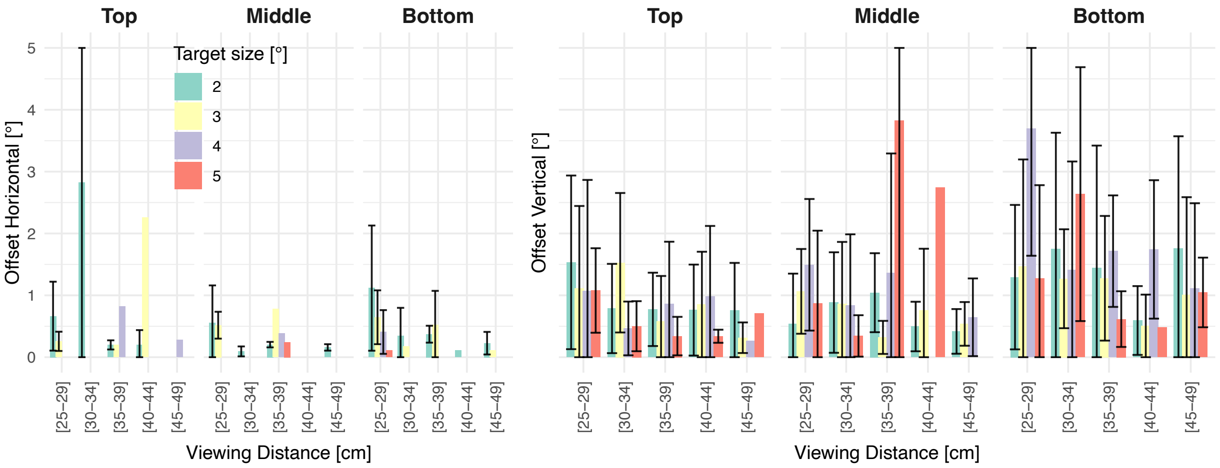

Figure 5: Mean offset errors for each viewing distance level, target size, and region with error bars representing the standard deviation. The plot only includes errors where the mean estimated gaze points fall outside the target area, indicating an offset error. Errors were generally higher in the vertical direction compared to the horizontal direction.

Observation 3: Tracking accuracy is significantly worse in the vertical direction compared to the horizontal one.

4.2 Tracking Precision

We conducted a Wilcoxon signed-rank test to evaluate the difference between the horizontal and vertical directions on the collected data's precision (Standard deviation). The results indicated a statistically significant difference in the precision between the two directions ($V = 50,511, p < .0005$). This suggests that the precision of the collected gaze data was worse in the vertical direction ($Med = .484°, Q1 = .293°, Q3 = .792°$) compared to the horizontal direction ($Med = .173°, Q1 = .104°, Q3 = .271°$) (see Figure 4).

Observation 4: In the horizontal direction, targets of size 3° maintain tracking accuracy above 80% across all viewing distances.

4.3 Tracking Offset Error

Figure 5 shows the offset error for each target size and region at each level of viewing distance in both the horizontal and vertical directions. Errors were generally higher in the vertical direction ($M = 1.07°, SD = 1.24$) compared to the horizontal direction ($M = 0.561°, SD = 0.793$). In the vertical direction, the target located at the bottom region of the screen resulted in higher errors ($M = 1.44°, SD = 1.51$) compared to targets located at the top ($M = 0.865°, SD = 0.955$) and middle regions ($M = 0.832°, SD = 1.03$). Figure 6 visualises how the error increase in the vertical direction and how targets located in the top and bottom regions of the phone screen caused a greater offset towards the center, suggesting a tendency for participants' gaze to drift towards the center when aiming at these edge targets.

Observation 5: Precision of the collected gaze data was worse in the vertical direction compared to the horizontal one. Similarly, the offset errors were higher in the vertical direction.

4.4 Effect of Viewing Distance on Tracking Accuracy

Results from Table 1 show that tracking accuracy decreases as the distance decreases. To explore whether such results are impacted by an intermediate variable, which is the computed target size that is calculated based on the formula that takes into account the target size in visual angle and the viewing distance (see Formula 1), we conducted a causal mediation analysis [29] to examine whether the effect of *Viewing Distance* on the tracking accuracy is mediated by the computed target sizes (the intermediate variable). This analysis helps to understand the underlying mechanism by which *Viewing Distance* influenced the tracking accuracy. We employed a quasi-Bayesian approach to estimate the intermediate variable effect, using nonparametric bootstrap confidence intervals with the percentile method [54]. The analysis revealed that the Average Causal Mediation Effect (ACME) was significant, with an estimate of 0.9738 ($95\%CI[0.8245, 1.12], p < .0005$), indicating that the computed target sizes significantly played as an intermediate variable in the relationship between the *Viewing Distance* and tracking accuracy. The Average Direct Effect (ADE) was not significant, with an estimate of 0.0255 ($95\%CI[-0.2628, 0.30], p = 0.86$), suggesting that the direct effect of *Viewing Distance* on the tracking accuracy is negligible when controlling for the intermediate variable. The Total Effect of *Viewing Distance* on the tracking accuracy was significant ($Estimate = 0.9993, 95\%CI[0.7231, 1.24], p < .0005$). These findings

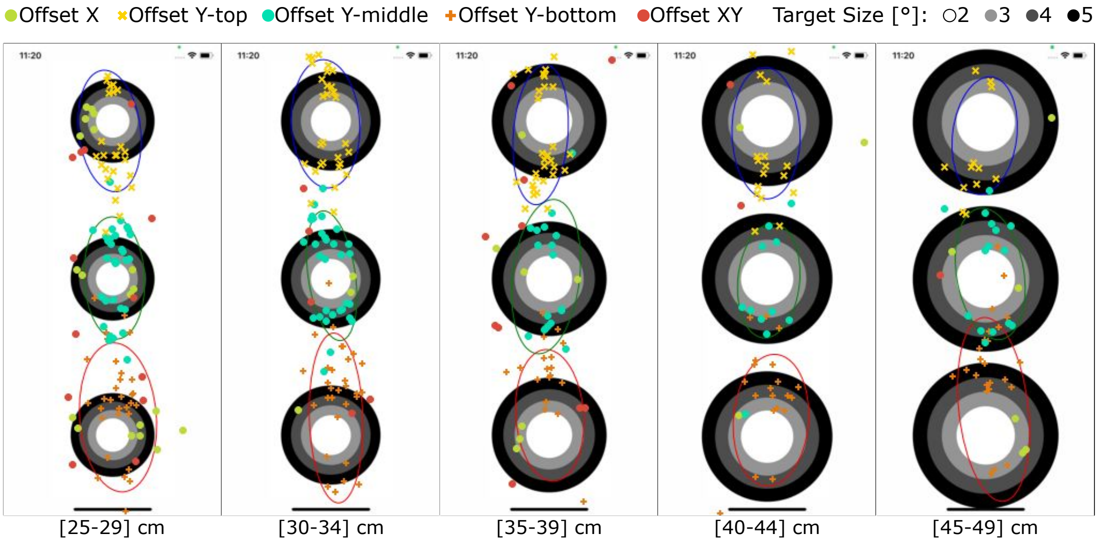

Figure 6: Offset error across various target sizes for all viewing distances. Errors are displayed in different colours depending on whether the error is only in horizontal, vertical, or both directions. Each target size is depicted by a unique colour gradient ranging from black to white. Each screen displays the offset across three distinct regions: top, middle, and bottom. Covariance ellipses depict the contours of the 2D Gaussian, fitted to the gaze points associated with all fixations on the same target. The figures show that errors were generally higher in the vertical direction compared to the horizontal direction.

suggest that the majority of the effect of the *Viewing Distance* on the accuracy is coming through the computed target sizes, highlighting their significant impacts on the results.

Observation 6: Targets located at top and bottom regions of the phone resulted in an increased vertical offset error, suggesting a tendency to drift toward the centre of the screen.

4.5 Impact of Target Size and Target Tracking Area on Tracking Accuracy

To study the effect of target size and target tracking area on tracking accuracy, we tested the impact of expanding the tracking areas, namely 1) uniform expansion, 2) horizontal expansion, and 3) vertical expansion, and compared it to targets with equal visual and tracking areas. We excluded the target with 5° due to screen size.

Observation 7: The on-screen target dimensions, computed in units such as centimetres, affect tracking accuracy when adjusted to preserve the same visual angle relative to different viewing distances.

4.5.1 Uniform Tracking Area Expansion. We consistently adjusted the target dimensions at each viewing distance, applying the same modifications to both width and height to ensure equal expansion in both directions. For visual targets of size 2°x2°, accuracy generally improved when paired with a tracking area of 3°x3° ($M = 70.6\%, SD = 35.4$) compared to targets with a visual and tracking area of 3°x3° ($M = 67.1\%, SD = 35.6$), and for visual targets of size 2°x2° with tracking area of 4°x4° ($M = 80.5\%, SD = 32.2$) compared to targets with a visual and tracking area of 4°x4° ($M = 77.8\%, SD = 32.0$). Similarly, targets with a visual size of 3°x3° combined with tracking area of 4°x4° ($M = 80.6\%, SD = 29.8$) or 5°x5° ($M = 86.9\%, SD = 25.6$) exhibited increased tracking accuracy compared to a target with both visual and tracking area of 4°x4° ($M = 77.8\%, SD = 32.0$) or 5°x5° ($M = 86.0\%, SD = 25.8$) respectively. However, visual targets of size 4°x4° with a tracking area of 5°x5° ($M = 85.8\%, SD = 26.6$) did not exhibit an increase in accuracy compared to targets with a visual and tracking area of 5°x5°.

For targets with visuals of 2°x2° and 3°x3°, with tracking areas of 3°x3°, repeated measure ANOVA test did not reveal a significant main effect of *Target Size*, $F_{1,\,23}=3.29$, $p > .05$, nor there was an interaction effect. Similarly, the test did not reveal a significant main effect of target size on the tracking accuracy, for targets with visuals 2°x2°, 3°x3°, and 4°x4° with tracking areas of 4°x4°, $F_{1.93,\,44.44}=1.18$, $p > .05$, nor there were significant differences among targets with visual sizes of 2°x2°, 3°x3°, 4°x4°, and 5°x5° with tracking areas of 5°x5°, as the test did not reveal significant main effect of target size, $F_{2.74,\,62.92}=0.24$, $p > .05$. Overall, while not significant, smaller visual targets benefit from increasing the tracking areas compared to bigger targets with uniform visual and tracking areas.

4.5.2 Horizontal Tracking Area Expansion. We explored the impact of expanding the width of the tracking area on the tracking accuracy in the horizontal direction for each target size used in the experiment. The tracking area widths were determined based on

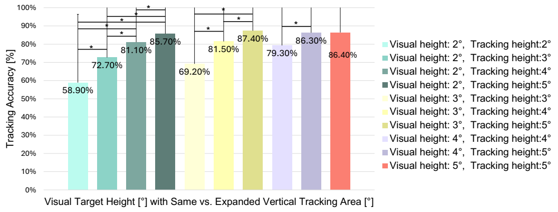

Figure 7: Mean tracking accuracy in the vertical direction for different visual targets compared to the same targets when their tracking area is expanded vertically. For each target size, significant differences were found between all pairs, suggesting that the vertical direction is highly impacted by expanding the tracking area in the vertical direction. The error bar represents the standard deviation.

the target sizes used in the experiment. We only adjusted the width when necessary, ensuring that the width remained greater or equal to the height. For each level of *Target Size*, we performed repeated measures ANOVA.

Observation 8: Expanding the tracking area can improve accuracy, with the effect being more pronounced in the vertical direction than in the horizontal direction.

For targets of visual width 2°, there was a significant interaction effect between *Viewing Distance* and *Target Width* on the accuracy, $F_{3.32,\ 76.27}=3.48$, $p < .05$. The pairwise comparison revealed significantly reduced accuracy with targets of tracking width 2° ($M = 63.9\%, SD = 40.4$) compared to tracking width 3° ($M = 78.8\%, SD = 36.4, p < .05$), 4° ($M = 83.4\%, SD = 35.5, p < .01$), and 5° ($M = 85.3\%, SD = 34.9, p < .01$) when viewed at a distance of [25-29] cm. No significant differences were found at other distances.

While we found a significant interaction between *Viewing Distance* and *Target Width* on the accuracy, $F_{2.96,\ 68.04}=2.87$, $p < .05$, for targets of visual width 3°, post hoc analysis did not reveal significant differences between pairs.

For targets of visual width 4°, there was a significant main effect of *Target Width* on the tracking accuracy, $F_{1,\ 23}=9.75$, $p < .01$. Targets of tracking width 4° ($M = 97.0\%, SD = 14.2$) performed significantly lower ($p < .01$) compared to targets with tracking width 5° ($M = 98.3\%, SD = 11.3$).

Our findings show that increasing the width of the tracking area significantly enhances the horizontal tracking accuracy for targets of 2° at the closest distance, e.g. [25-29], and also for relatively larger targets, e.g. 4°, regardless of the distance.

4.5.3 Vertical Tracking Area Expansion. Similar to studying the effect of changing the tracking area width, we also experimented with changing its height. This was inspired by our results from sections 4.1, 4.2, and 4.3, which showed that the gaze data from the vertical direction impacted the tracking accuracy. The heights were determined based on the target sizes used in the experiment. We only adjusted the height when necessary, ensuring that the height remained greater or equal to the width. We excluded targets of 5° size when increasing the targets' heights, as this adjustment caused the tracking area to exceed the screen size. For each level of *Target Size*, we performed repeated measures ANOVA. Figure 7 shows the statistically significant pairs.

For targets of visual height 2°, there was a significant main effect of *Viewing Distance*, $F_{2.98,\ 68.49}=6.40$, $p < .001$, and *Target Height*, $F_{1.64,\ 37.6}=275.63$, $p < .001$ on the accuracy of the collected gaze data in the vertical direction. There was no interaction effect between all factors. Pairwise comparisons revealed statistically significant differences between all heights, where increasing the tracking area increased the tracking accuracy. Targets viewed at viewing distance of [40-44] cm ($M = 85.1\%, SD = 27$) also resulted in significantly improved accuracy ($p < .005$) compared to targets viewed at distances of [25-29] ($M = 61.8\%, SD = 38.5$) and [30-34] ($M = 68.3\%, SD = 37.8$) cm.

For targets of visual height 3°, there was a significant main effect of *Viewing Distance*, $F_{2.85,\ 65.62}=6.24$, $p < .005$, and *Target Height*, $F_{1.43,\ 32.81}=197.05$, $p < .001$ on the accuracy of the collected gaze data in the vertical direction. Targets at viewing distances of [40-44] ($M = 87.0\%, SD = 25.0$) and [45-49] cm ($M = 86.1\%, SD = 23.9$) resulted in significantly better tracking accuracy compared to targets viewed at the viewing distance of [25-29] cm ($M = 68.0\%, SD = 36.6$). For target heights, pairwise comparisons revealed statistically significant differences between all heights (see Figure 7).

For targets of visual height 4°, there was a significant main effect of *Target Height*, $F_{1,\ 23}=41.37$, $p < .001$ on the accuracy. pairwise comparisons revealed statistically significant difference ($p < .0001$) between targets of tracking height 4° ($M = 79.3\%, SD = 31.0$) and 5° ($M = 86.3\%, SD = 26.3$).

While the results showed a general increase in the vertical tracking accuracy when expanding the vertical tracking area to 4° and maintaining smaller visual targets in terms of height (see Figure 7), 2° ($M = 81.1\%, SD = 31.9$), 3° ($M = 81.5\%, SD = 29.0$), compared to 4° ($M = 79.3\%, SD = 31.0$), one-way ANOVA results indicated no significant differences between the groups $F_{2,\ 1077}=0.52$, $p > .05$.

4.6 Subjective Feedback

Participants responded on a Likert scale, where 1 represented "*Strongly disagree*" and 5 represented "*Strongly agree*", to whether they found fixating on the target tiring to their eyes. At the *Viewing Distance* of [25-29] cm, using ART, the statistical test revealed a significant main effect of *Target Size*, $F(3)= 4.588$, $p < .01$, and *Target Region*, $F(2)= 3.797$, $p < .05$, on the perceived eye fatigue by the participants, with no interaction effect. Pairwise comparison showed that participants perceived targets of size 2° as significantly more tiring to the eyes ($Med = 2, Q1 = 2, Q3 = 4$) compared to targets of size 4° ($Med = 2, Q1 = 1, Q3 = 3, p < .05$) and 5° ($Med = 2, Q1 = 1, Q3 = 3, p < .01$). Targets located at the bottom region were also perceived as significantly more tiring ($Med = 2, Q1 = 1, Q3 = 3.25, p < .05$) compared to targets placed in the middle region ($Med = 2, Q1 = 1, Q3 = 3, p < .05$). At the *Viewing Distance* of [30-34] cm, only a significant main effect of *Target Size* was found on the perceived tiredness to the eyes, $F(3)= 6.16$, $p < .0005$. Pairwise comparison showed that participants perceived targets of size 2° as significantly more tiring to their eyes ($Med = 2, Q1 = 2, Q3 = 3.25$) compared targets of size 4° ($Med = 2, Q1 = 1, Q3 = 3, p < .01$) and 5° ($Med = 2, Q1 = 1, Q3 =$

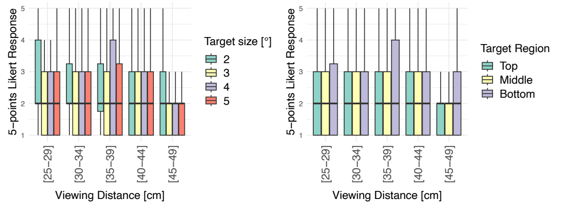

Figure 8: Participants' responses to whether they found fixating on the targets of various sizes at various viewing distances tiring to their eyes (1= Strongly disagree; 5= Strongly agree). The box represents the 25th, 50th, and 75th percentiles. Left: Responses based on target size. Right: Responses based on target region. Overall, participants perceived targets of size 2° as more tiring to their eyes than targets of size 4° and 5° at most viewing distance levels.

3, $p < .001$). At the *Viewing Distance* of [35-39], we found no significant main effect of *Target Size*, *Target Region*, no there was an interaction effect between them. While at the *Viewing Distance* of [40-44] cm, we found a significant main effect for *Target Size* on the perceived tiredness to the eyes, F(3)= 3.17, $p < .05$, no significant differences were found between pairs. The statistical test also revealed significant main effect of *Target Size*, F(3)= 5.66, $p < .001$, and *Target Region*, F(2)= 6.41, $p < .01$, on the perceived eye fatigue by the participants at the *Viewing Distance* of [45-49] cm. Pairwise comparison showed that participants perceived targets of size 2° as significantly more tiring to the eyes ($Med = 2, Q1 = 1, Q3 = 3$) during fixation compared to all other sizes of 3° ($p < .05$), 4° ($p < .01$), and 5° ($p < .01$) ($Med = 2, Q1 = 1, Q3 = 2$). Targets placed at the middle ($Med = 2, Q1 = 1, Q3 = 2, p < .05$) and bottom ($Med = 2, Q1 = 1, Q3 = 3, p < .01$) regions were perceived as significantly more tiring to the eyes compared to targets placed at top region ($Med = 2, Q1 = 1, Q3 = 2$).

5 DISCUSSION

We showed how tracking accuracy and precision can vary depending on the viewing distance and the target size in visual angle and how tracking in the horizontal direction is more accurate than tracking in the vertical direction. Our work underscores the importance of considering the variation in the distance between the user and the mobile phone when designing the target size, as the visual angle alone is not enough to determine that. In the following, we discuss our findings, provide recommendations for target size when designing gaze-enabled interfaces for mobile devices, and discuss opportunities for future work.

5.1 Increasing Target Height Compensates for Vertical Tracking Errors on Gaze Interfaces

While the results showed that targets of size 4° serve as an optimal sweet spot across all face-to-screen distances, maintaining a tracking accuracy of at least 70% with minimal exceptions, targets of size 3° maintain an accuracy higher than 80% in the horizontal direction. The results also showed that tracking data in the vertical direction resulted in lower precision and higher offset errors than horizontal data. Based on these findings, we suggest that targets' heights for gaze-enabled interfaces in mobile devices must be large enough to compensate for such tracking errors. Findings from our study are consistent with prior work on target size and tracking accuracy on desktops using dedicated eye trackers [22], which recommended designing tracking regions with greater height than width to enhance accuracy. Other studies examining target motion have found that tracking is more accurate in the horizontal direction compared to the vertical direction [11, 58]. Vertical saccades were reported as less accurate compared to horizontal ones [9, 10]. We hypothesise that such variation between horizontal and vertical directions is due to the horizontal-vertical anisotropy phenomenon affecting human visual perception and cognition, where humans tend to perceive horizontal and vertical dimensions differently [62]. Given that mobile device screens are typically longer vertically than horizontally, we hypothesise that this factor may exacerbate the issue, leading to larger Euclidean errors when estimating gaze position. Future work could investigate the role of mobile device orientation in portrait versus landscape in minimising the negative effect of the vertical direction on tracking accuracy to enhance the usability of gaze interaction with mobile devices.

Observation 9: The variation in tracking accuracy, precision, and offset errors between the horizontal and vertical directions may be due to the horizontal-vertical anisotropy phenomenon. We hypothesise that the greater vertical length of the mobile screen compared to its horizontal dimension further intensifies the issue.

5.2 Expanding Tracking Areas Improves Accuracy but Requires Overlap Management

The further analysis we conducted to expand the tracking area showed that expanding the tracking area generally enhances the tracking accuracy. However, such expansion in the horizontal direction did not result in a significant positive impact on the accuracy of tracking at most levels of *Target Size.* For the vertical direction, expanding the tracking region significantly improved tracking accuracy in all levels of *Target Size.* While the results align with prior work indicating an improved gaze-pointing performance when expanding the tracking area [41, 66], placing the targets next to each other may negatively impact the performance as the invisible tracking area of adjacent targets may overlap. While maintaining enough room for tracking areas may help reduce the issue's impact and utilise such space for non-interactive objects, given the limited mobile screen's real estate where spaces need to be utilised, future work could explore the overlapped tracking areas and find ways to reduce their impact. For example, dynamically adjusting the tracking area based on how far the targets are from each other or shrinking the adjacent tracking area based on the user gaze duration over a target can be a direction. Another direction is to allow the systems to detect the user's gaze path and stick to the most likely target based on how the user behaves during the interaction. Using machine learning, user intent can be detected alongside the gaze path to measure which target users are moving their gaze toward and select them.

Table 3: Recommended target sizes for gaze-enabled interfaces. Given the distance, we picked the minimum accuracy of the tracking to be a minimum of 70%. For target size, The first four rows display the recommended target size in terms of width and height in visual angles and their corresponding accuracy based on horizontal, vertical, or both directions. The following rows show the recommended target size for each direction individually. Based on our findings and prior work [22, 47], we suggest avoiding placing targets towards the edge of the screen as it degrades tracking accuracy.

Minimum Visual Target width [°] x height [°]	Expected Accuracy			Minimum Viewing Distance
	Horizontal-Vertical	Horiztonal	Vertical	
2 x 4	-	> 80%	> 70%	30 cm
3 x 4	-	> 80%	> 70%	25 cm
3 x 3	> 70%	> 90%	> 70%	40 cm
4 x 4	> 70%	> 90%	> 70%	25 cm
Visual Target Width [°]	**Minimum Tracking width [°]**	**Expected Accuracy [Horizontal]**		**Minimum Viewing Distance**
2	**3**	> 80%		25 cm
3	**3**	> 80%		25 cm
4	**4**	> 90%		25 cm
Visual Target Height [°]	**Minimum Tracking Height [°]**	**Expected Accuracy [Vertical]**		**Minimum Viewing Distance**
2	4	> 80%		35 cm
3	4	> 80%		30 cm
4	5	> 80%		25 cm

5.3 Recommendations for Target Sizes and Design Guidelines for Gaze-enabled Interfaces on Mobile Devices

Based on findings from our experiment, we present table 3, which shows our recommendations for minimum target sizes when designing for gaze selection on mobile devices, considering the viewing distance between the users and the screen. The table suggests target sizes in visual angles at various viewing distances based on both Horizontal-vertical directions or based on each direction individually. Based on the findings, we provide the following guidelines when designing for target selection in gaze-enabled interfaces:

[1] Design targets to be taller vertically than they are wide horizontally to compensate for tracking inaccuracies.
[2] Designing targets with larger tracking areas than their visual stimuli could improve tracking precision and accuracy, particularly in the vertical direction.
[3] We recommend designing targets with smaller visual angles, such as 3°, for further distances between users and phones. This will help maintain good tracking accuracy while conserving screen space. While increasing the target sizes improves accuracy, there is a trade-off, as this may limit the number of targets you can place on the screen.
[4] Use the recommended target sizes in Table 3 based on typical viewing distances for each phone application, as well as the user's context.
[5] Avoid placing targets at the top and bottom edges of the screen. If deemed necessary, we recommend increasing their vertical dimension to account for errors in such regions.

Although the recommendations and design guidelines are influenced by our setups, considering the varying distances at which users hold their devices while interacting with them [12, 27, 34], our findings pave the way for robust and more usable gaze selection on mobile devices and open new opportunities. Future research can build on this work to move mobile gaze input beyond the lab and assess the usability in mobile settings, including shaky environments and mobile contexts like sitting or walking. Developers can mitigate the negative impact of distance variations on gaze input by leveraging device sensors to detect context and viewing distance, enabling dynamic adjustment of target sizes. For example, UI element sizes can be adapted to maintain a consistent visual angle, accounting for distance variations in different mobile devices holding postures [27]. Adjustments to target sizes can also be tailored to specific applications. Depending on the application category, some applications are often used at closer distances, while others are used at farther distances. Our results are based on the phone positioned with a fixed angle at 60° [27]. However, we expect angle changes when using the phone in the mobile context. Such changes might result in the face being not visible or the eyes not being tracked by the camera. Guiding users to the best holding posture could be a way to mitigate such impact [33]. Guiding methods that were proposed for guiding users in front of public displays may be promising if adapted to mobile devices for this purpose [2, 70].

6 CONCLUSION

In this work, we explored various target sizes, regions, and viewing distances to optimise target selection for dwell-based interactions on mobile devices. Our findings showed that a visual angle of 4° serves as an optimal target size across all user distances. We also discovered that increasing the tracking area relative to the visual target size improves tracking accuracy, but only for smaller targets, such as those around 2°. Additionally, we provided design guidelines for researchers and designers to effectively leverage gaze input on mobile devices.

ACKNOWLEDGMENTS

This work has been funded by the Islamic University of Madinah.

REFERENCES

[1] Yasmeen Abdrabou, Ahmed Shams, Mohamed Omar Mantawy, Anam Ahmad Khan, Mohamed Khamis, Florian Alt, and Yomna Abdelrahman. 2021. GazeMeter: Exploring the Usage of Gaze Behaviour to Enhance Password Assessments. In *ACM Symposium on Eye Tracking Research and Applications* (Virtual Event, Germany) *(ETRA '21 Full Papers)*. Association for Computing Machinery, New York, NY, USA, Article 9, 12 pages. https://doi.org/10.1145/3448017.3457384

[2] Florian Alt, Andreas Bulling, Gino Gravanis, and Daniel Buschek. 2015. GravitySpot: Guiding Users in Front of Public Displays Using On-Screen Visual Cues. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Charlotte, NC, USA) *(UIST '15)*. Association for Computing Machinery, New York, NY, USA, 47–56. https://doi.org/10.1145/2807442.2807490

[3] Apple Inc. 2024. App icons. https://developer.apple.com/design/human-interface-guidelines/foundations/app-icons/ Accessed: 15 August 2024.

[4] Apple Inc. 2024. Control iPhone with the movement of your eyes. https://support.apple.com/en-gb/guide/iphone/iph66057d0f6/ios Accessed: 11 November 2024.

[5] Michael Ashmore, Andrew T. Duchowski, and Garth Shoemaker. 2005. Efficient eye pointing with a fisheye lens. In *Proceedings of Graphics Interface 2005* (Victoria, British Columbia) *(GI '05)*. Canadian Human-Computer Communications Society, Waterloo, CAN, 203–210.

[6] Irfan Ullah Baig. 2023. Comparing the Aesthetics of Rounded and Rectangular Buttons. Webpage. Retrieved September 10, 2024 from https://medium.com/@IrfanUlahBaig/comparing-the-aesthetics-of-rounded-and-rectangular-buttons-6294cae40061

[7] Darrell S. Best and Andrew T. Duchowski. 2016. A rotary dial for gaze-based PIN entry. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications* (Charleston, South Carolina) *(ETRA '16)*. Association for Computing Machinery, New York, NY, USA, 69–76. https://doi.org/10.1145/2857491.2857527

[8] WX Chen, XY Cui, J Zheng, JM Zhang, S Chen, and YD Yao. 2019. Gaze Gestures and Their Applications in human-computer interaction with a head-mounted display. *arXiv preprint arXiv:1910.07428* (2019).

[9] Han Collewijn, Casper J Erkelens, and Robert M Steinman. 1988. Binocular co-ordination of human horizontal saccadic eye movements. *The Journal of physiology* 404, 1 (1988), 157–182.

[10] Han Collewijn, Casper J Erkelens, and Robert M Steinman. 1988. Binocular co-ordination of human vertical saccadic eye movements. *The Journal of physiology* 404, 1 (1988), 183–197.

[11] Frederic R. Danion, James Mathew, Niels Gouirand, and Eli Brenner. 2021. More precise tracking of horizontal than vertical target motion with both the eyes and hand. *Cortex* 134 (2021), 30–42. https://doi.org/10.1016/j.cortex.2020.10.001

[12] Anind K. Dey, Katarzyna Wac, Denzil Ferreira, Kevin Tassini, Jin-Hyuk Hong, and Julian Ramos. 2011. Getting closer: an empirical investigation of the proximity of user to their smart phones. In *Proceedings of the 13th International Conference on Ubiquitous Computing* (Beijing, China) *(UbiComp '11)*. Association for Computing Machinery, New York, NY, USA, 163–172. https://doi.org/10.1145/2030112.2030135

[13] Heiko Drewes, Alexander De Luca, and Albrecht Schmidt. 2007. Eye-Gaze Interaction for Mobile Phones. In *Proceedings of the 4th International Conference on Mobile Technology, Applications, and Systems and the 1st International Symposium on Computer Human Interaction in Mobile Technology* (Singapore) *(Mobility '07)*. Association for Computing Machinery, New York, NY, USA, 364–371. https://doi.org/10.1145/1378063.1378122

[14] Heiko Drewes and Albrecht Schmidt. 2007. Interacting with the Computer Using Gaze Gestures. In *Proceedings of the 11th IFIP TC 13 International Conference on Human-Computer Interaction - Volume Part II* (Rio de Janeiro, Brazil) *(INTERACT'07)*. Springer-Verlag, Berlin, Heidelberg, 475–488.

[15] Heiko Drewes and Albrecht Schmidt. 2009. The MAGIC Touch: Combining MAGIC-Pointing with a Touch-Sensitive Mouse. In *Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: Part II* (Uppsala, Sweden) *(INTERACT '09)*. Springer-Verlag, Berlin, Heidelberg, 415–428. https://doi.org/10.1007/978-3-642-03658-3_46

[16] Andrew Duchowski, Eric Medlin, Nathan Cournia, Hunter Murphy, Anand Gramopadhye, Santosh Nair, Jeenal Vorah, and Brian Melloy. 2002. 3-D eye movement analysis. *Behavior Research Methods, Instruments, & Computers* 34 (2002), 573–591.

[17] Morten Lund Dybdal, Javier San Agustin, and John Paulin Hansen. 2012. Gaze Input for Mobile Devices by Dwell and Gestures. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (Santa Barbara, California) *(ETRA '12)*. Association for Computing Machinery, New York, NY, USA, 225–228. https://doi.org/10.1145/2168556.2168601

[18] Augusto Esteves, Elizabeth Bouquet, Ken Pfeuffer, and Florian Alt. 2022. One-handed Input for Mobile Devices via Motion Matching and Orbits Controls. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 2, Article 51 (jul 2022), 24 pages. https://doi.org/10.1145/3534624

[19] Augusto Esteves, Yonghwan Shin, and Ian Oakley. 2020. Comparing selection mechanisms for gaze input techniques in head-mounted displays. *International Journal of Human-Computer Studies* 139 (2020), 102414.

[20] Augusto Esteves, Eduardo Velloso, Andreas Bulling, and Hans Gellersen. 2015. Orbits: Gaze Interaction for Smart Watches Using Smooth Pursuit Eye Movements. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Charlotte, NC, USA) *(UIST '15)*. Association for Computing Machinery, New York, NY, USA, 457–466. https://doi.org/10.1145/2807442.2807499

[21] Martha J Farah, Michael J Soso, and Richard M Dasheiff. 1992. Visual angle of the mind's eye before and after unilateral occipital lobectomy. *Journal of Experimental Psychology: Human Perception and Performance* 18, 1 (1992), 241.

[22] Anna Maria Feit, Shane Williams, Arturo Toledo, Ann Paradiso, Harish Kulkarni, Shaun Kane, and Meredith Ringel Morris. 2017. Toward Everyday Gaze Input: Accuracy and Precision of Eye Tracking and Implications for Design. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 1118–1130. https://doi.org/10.1145/3025453.3025599

[23] Misahael Fernandez, Florian Mathis, and Mohamed Khamis. 2020. *GazeWheels: Comparing Dwell-Time Feedback and Methods for Gaze Input.* Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3419249.3420122

[24] Li-Chen Fu and Robin Fischer. 2023. Collaboration Assistance Through Object Based User Intent Detection Using Gaze Data. In *Proceedings of the 2023 Symposium on Eye Tracking Research and Applications* (Tubingen, Germany) *(ETRA '23)*. Association for Computing Machinery, New York, NY, USA, Article 35, 2 pages. https://doi.org/10.1145/3588015.3590130

[25] Dan Witzner Hansen and Qiang Ji. 2010. In the Eye of the Beholder: A Survey of Models for Eyes and Gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, 3 (2010), 478–500. https://doi.org/10.1109/TPAMI.2009.30

[26] Henna Heikkilä and Kari-Jouko Räihä. 2012. Simple Gaze Gestures and the Closure of the Eyes as an Interaction Technique. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (Santa Barbara, California) *(ETRA '12)*. Association for Computing Machinery, New York, NY, USA, 147–154. https://doi.org/10.1145/2168556.2168579

[27] Michael Xuelin Huang, Jiajia Li, Grace Ngai, and Hong Va Leong. 2017. ScreenGlint: Practical, In-Situ Gaze Estimation on Smartphones. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2546–2557. https://doi.org/10.1145/3025453.3025794

[28] Qiong Huang, Ashok Veeraraghavan, and Ashutosh Sabharwal. 2017. TabletGaze: dataset and analysis for unconstrained appearance-based gaze estimation in mobile tablets. *Machine Vision and Applications* 28, 5 (2017), 445–461.

[29] Kosuke Imai, Luke Keele, and Dustin Tingley. 2010. A general approach to causal mediation analysis. *Psychological methods* 15, 4 (2010), 309.

[30] Mahdi Tavakoli Isao Sakamaki, Kim Adams and Sandra A. Wiebe. 2021. Effect of feedback and target size on eye gaze accuracy in an off-screen task. *Disability and Rehabilitation: Assistive Technology* 16, 7 (2021), 769–779. https://doi.org/10.1080/17483107.2020.1729874 arXiv:https://doi.org/10.1080/17483107.2020.1729874 PMID: 32100583.

[31] Robert J. K. Jacob. 1990. What You Look at is What You Get: Eye Movement-Based Interaction Techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, USA) *(CHI '90)*. Association for Computing Machinery, New York, NY, USA, 11–18. https://doi.org/10.1145/97243.97246

[32] Robert J. K. Jacob. 1991. The Use of Eye Movements in Human-Computer Interaction Techniques: What You Look at is What You Get. *ACM Trans. Inf. Syst.* 9, 2 (apr 1991), 152–169. https://doi.org/10.1145/123078.128728

[33] Mohamed Khamis, Florian Alt, and Andreas Bulling. 2018. The Past, Present, and Future of Gaze-Enabled Handheld Mobile Devices: Survey and Lessons Learned. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Barcelona, Spain) *(MobileHCI '18)*. Association for Computing Machinery, New York, NY, USA, Article 38, 17 pages. https://doi.org/10.1145/3229434.3229452

[34] Mohamed Khamis, Anita Baier, Niels Henze, Florian Alt, and Andreas Bulling. 2018. Understanding Face and Eye Visibility in Front-Facing Cameras of Smartphones Used in the Wild. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173854

[35] Andy Kong, Karan Ahuja, Mayank Goel, and Chris Harrison. 2021. EyeMU Interactions: Gaze+ IMU Gestures on Mobile Devices. In *Proceedings of the 2021 International Conference on Multimodal Interaction*. 577–585.

[36] Manu Kumar, Andreas Paepcke, and Terry Winograd. 2007. EyePoint: practical pointing and selection using gaze and keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '07)*. Association for Computing Machinery, New York, NY, USA, 421–430.

https://doi.org/10.1145/1240624.1240692

[37] Chris Lankford. 2000. Effective eye-gaze input into Windows. In *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications* (Palm Beach Gardens, Florida, USA) *(ETRA '00)*. Association for Computing Machinery, New York, NY, USA, 23–27. https://doi.org/10.1145/355017.355021

[38] Yaxiong Lei, Yuheng Wang, Tyler Caslin, Alexander Wisowaty, Xu Zhu, Mohamed Khamis, and Juan Ye. 2023. DynamicRead: Exploring Robust Gaze Interaction Methods for Reading on Handheld Mobile Devices under Dynamic Conditions. *Proceedings of the ACM on Human-Computer Interaction* 7, ETRA (2023), 1–17.

[39] Christof Lutteroth, Moiz Penkar, and Gerald Weber. 2015. Gaze vs. Mouse: A Fast and Accurate Gaze-Only Click Alternative. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Charlotte, NC, USA) *(UIST '15)*. Association for Computing Machinery, New York, NY, USA, 385–394. https://doi.org/10.1145/2807442.2807461

[40] Päivi Majaranta and Andreas Bulling. 2014. Eye tracking and eye-based human–computer interaction. In *Advances in physiological computing*. Springer, 39–65.

[41] Darius Miniotas, Oleg Špakov, and I. Scott MacKenzie. 2004. Eye gaze interaction with expanding targets. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria) *(CHI EA '04)*. Association for Computing Machinery, New York, NY, USA, 1255–1258. https://doi.org/10.1145/985921.986037

[42] Mahya Mirbagheri and Tom Chau. 2024. Optimising virtual object position for efficient eye-gaze interaction in Hololens2. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* 12, 1 (2024), 2337765. https://doi.org/10.1080/21681163.2024.2337765 arXiv:https://doi.org/10.1080/21681163.2024.2337765

[43] Emilie Mollenbach, John Paulin Hansen, Martin Lillholm, and Alastair G. Gale. 2009. Single Stroke Gaze Gestures. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI EA '09)*. Association for Computing Machinery, New York, NY, USA, 4555–4560. https://doi.org/10.1145/1520340.1520699

[44] Emilie Møllenbach, Martin Lillholm, Alastair Gail, and John Paulin Hansen. 2010. Single Gaze Gestures. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (Austin, Texas) *(ETRA '10)*. Association for Computing Machinery, New York, NY, USA, 177–180. https://doi.org/10.1145/1743666.1743710

[45] Akito Monden, Ken-ichi Matsumoto, and Masatake Yamato. 2005. Evaluation of gaze-added target selection methods suitable for general GUIs. *International journal of computer applications in technology* 24, 1 (2005), 17–24.

[46] Emilie Møllenbach, John Paulin Hansen, and Martin Lillholm. 2013. Eye Movements in Gaze Interaction. *Journal of Eye Movement Research* 6, 2 (May 2013). https://doi.org/10.16910/jemr.6.2.1

[47] Omar Namnakani, Yasmeen Abdrabou, Jonathan Grizou, Augusto Esteves, and Mohamed Khamis. 2023. Comparing Dwell time, Pursuits and Gaze Gestures for Gaze Interaction on Handheld Mobile Devices. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3544548.3580871

[48] Ya-feng Niu, Yue Gao, Ya-ting Zhang, Cheng-qi Xue, and Li-xin Yang. 2019. Improving eye–computer interaction interface design: Ergonomic investigations of the optimum target size and gaze-triggering dwell time. *Journal of Eye Movement Research* 12, 3 (2019).

[49] Abdul Moiz Penkar, Christof Lutteroth, and Gerald Weber. 2012. Designing for the Eye: Design Parameters for Dwell in Gaze Interaction. In *Proceedings of the 24th Australian Computer-Human Interaction Conference* (Melbourne, Australia) *(OzCHI '12)*. Association for Computing Machinery, New York, NY, USA, 479–488. https://doi.org/10.1145/2414536.2414609

[50] Ken Pfeuffer, Jason Alexander, Ming Ki Chong, and Hans Gellersen. 2014. Gaze-Touch: Combining Gaze with Multi-Touch for Interaction on the Same Surface. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST '14)*. Association for Computing Machinery, New York, NY, USA, 509–518. https://doi.org/10.1145/2642918.2647397

[51] Ken Pfeuffer, Jason Alexander, Ming Ki Chong, Yanxia Zhang, and Hans Gellersen. 2015. Gaze-Shifting: Direct-Indirect Input with Pen and Touch Modulated by Gaze. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Charlotte, NC, USA) *(UIST '15)*. Association for Computing Machinery, New York, NY, USA, 373–383. https://doi.org/10.1145/2807442.2807460

[52] Ken Pfeuffer and Hans Gellersen. 2016. Gaze and Touch Interaction on Tablets. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 301–311. https://doi.org/10.1145/2984511.2984514

[53] Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, and Hans Gellersen. 2017. Gaze + Pinch Interaction in Virtual Reality. In *Proceedings of the 5th Symposium on Spatial User Interaction* (Brighton, United Kingdom) *(SUI '17)*. Association for Computing Machinery, New York, NY, USA, 99–108. https://doi.org/10.1145/3131277.3132180

[54] Kristopher J Preacher and Andrew F Hayes. 2004. SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior research methods, instruments, & computers* 36 (2004), 717–731.

[55] Vijay Rajanna, Adil Hamid Malla, Rahul Ashok Bhagat, and Tracy Hammond. 2018. DyGazePass: A gaze gesture-based dynamic authentication system to counter shoulder surfing and video analysis attacks. In *2018 IEEE 4th International Conference on Identity, Security, and Behavior Analysis (ISBA)*. 1–8. https://doi.org/10.1109/ISBA.2018.8311458

[56] Argenis Ramirez Ramirez Gomez, Christopher Clarke, Ludwig Sidenmark, and Hans Gellersen. 2021. Gaze+Hold: Eyes-only Direct Manipulation with Continuous Gaze Modulated by Closure of One Eye. In *ACM Symposium on Eye Tracking Research and Applications* (Virtual Event, Germany) *(ETRA '21 Full Papers)*. Association for Computing Machinery, New York, NY, USA, Article 10, 12 pages. https://doi.org/10.1145/3448017.3457381

[57] Sheikh Rivu, Yasmeen Abdrabou, Thomas Mayer, Ken Pfeuffer, and Florian Alt. 2019. GazeButton: Enhancing Buttons with Eye Gaze Interactions. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications* (Denver, Colorado) *(ETRA '19)*. Association for Computing Machinery, New York, NY, USA, Article 73, 7 pages. https://doi.org/10.1145/3317956.3318154

[58] KLAUS G. ROTTACH, ARI Z. ZIVOTOFSKY, VALLABH E. DAS, LEA AVERBUCH-HELLER, ALFRED O. DISCENNA, ANUCHIT POONYATHALANG, and R.JOHN LEIGH. 1996. Comparison of Horizontal, Vertical and Diagonal Smooth Pursuit Eye Movements in Normal Human Subjects. *Vision Research* 36, 14 (1996), 2189–2195. https://doi.org/10.1016/0042-6989(95)00302-9

[59] Dario D. Salvucci and John R. Anderson. 2000. Intelligent gaze-added interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (The Hague, The Netherlands) *(CHI '00)*. Association for Computing Machinery, New York, NY, USA, 273–280. https://doi.org/10.1145/332040.332444

[60] Linda E. Sibert and Robert J. K. Jacob. 2000. Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (The Hague, The Netherlands) *(CHI '00)*. Association for Computing Machinery, New York, NY, USA, 281–288. https://doi.org/10.1145/332040.332445

[61] Henrik Skovsgaard, Julio C. Mateo, John M. Flach, and John Paulin Hansen. 2010. Small-target selection with gaze alone. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (Austin, Texas) *(ETRA '10)*. Association for Computing Machinery, New York, NY, USA, 145–148. https://doi.org/10.1145/1743666.1743702

[62] Daniel T. Smith. 2022. A horizontal–vertical anisotropy in spatial short-term memory. *Visual Cognition* 30, 4 (2022), 245–253. https://doi.org/10.1080/13506285.2022.2042446 arXiv:https://doi.org/10.1080/13506285.2022.2042446

[63] Sophie Stellmach and Raimund Dachselt. 2012. Look & touch: gaze-supported target acquisition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2981–2990. https://doi.org/10.1145/2207676.2208709

[64] Mélodie Vidal, Andreas Bulling, and Hans Gellersen. 2013. Pursuits: Spontaneous Interaction with Displays Based on Smooth Pursuit Eye Movement and Moving Targets. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Zurich, Switzerland) *(UbiComp '13)*. Association for Computing Machinery, New York, NY, USA, 439–448. https://doi.org/10.1145/2493432.2493477

[65] Mélodie Vidal, Ken Pfeuffer, Andreas Bulling, and Hans W. Gellersen. 2013. Pursuits: Eye-Based Interaction with Moving Targets. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 3147–3150. https://doi.org/10.1145/2468356.2479632

[66] Qi-Jun Wang, Xiao-Xing Ma, Yi-Ni Lu, Du-Ming Wang, and Yu-Hao Sun. 2024. The impact of visual and motor space size on gaze-based target selection. *Australian Journal of Psychology* 76, 1 (2024), 2309384. https://doi.org/10.1080/00049530.2024.2309384 arXiv:https://doi.org/10.1080/00049530.2024.2309384

[67] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[68] Wang Yi-yan, Cheng Wen, Tian Jing-ze, Huang Wei-chi, Niu Ya-feng, Yang Wen-jun, and Xue Cheng-qi. 2024. Design recommendations of target size and tracking speed under circular and square trajectories for smooth pursuit with Euclidean algorithm in eye-control system. *Displays* 81 (2024), 102608. https://doi.org/10.1016/j.displa.2023.102608

[69] Xinyong Zhang, Xiangshi Ren, and Hongbin Zha. 2008. Improving eye cursor's stability for eye pointing tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 525–534. https://doi.org/10.1145/1357054.1357139

[70] Yanxia Zhang, Jörg Müller, Ming Ki Chong, Andreas Bulling, and Hans Gellersen. 2014. GazeHorizon: enabling passers-by to interact with public displays by gaze. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Seattle, Washington) *(UbiComp '14)*. Association for Computing Machinery, New York, NY, USA, 559–563. https://doi.org/10.1145/2632048.2636071

[71] Hui Zheng and Vivian Genaro Motti. 2018. Assisting Students with Intellectual and Developmental Disabilities in Inclusive Education with Smartwatches. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173924