Performance Analysis of Saccades for Primary and Confirmatory Target Selection

Aunnoy K Mutasim
amutasim@sfu.ca
School of Interactive Arts & Technology (SIAT)
Simon Fraser University
Vancouver, BC, Canada

Anil Ufuk Batmaz
ufuk.batmaz@concordia.ca
Computer Science & Software Engineering Department
Concordia University
Montreal, QC, Canada

Moaaz Hudhud Mughrabi
moaaz.mazen@stu.khas.edu.tr
Mechatronics Engineering Department
Kadir Has University
Istanbul, Turkey

Wolfgang Stuerzlinger
w.s@sfu.ca
School of Interactive Arts & Technology (SIAT)
Simon Fraser University
Vancouver, BC, Canada

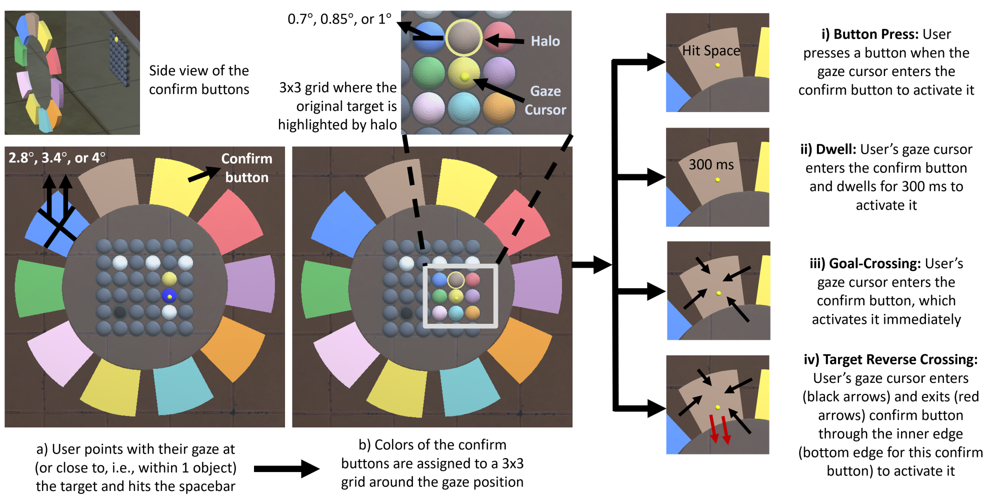


Figure 1: Actigaze with *Button Press* (ABP), *Dwell* (ADw), *Goal-Crossing* (AGC), or *Target Reverse Crossing* (ARC).

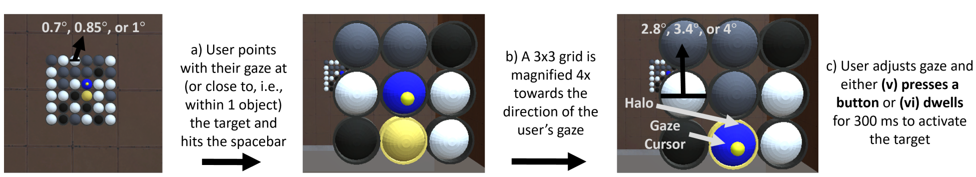


Figure 2: *4x Magnification* with *Button Press* (4xMBP) or *Dwell* (4xMDw). Viewpoint slightly offset to the left for illustration.

VRST '22, November 29-December 1, 2022, Tsukuba, Japan

ACM ISBN 978-1-4503-9889-3/22/11…$15.00
https://doi.org/10.1145/3562939.3565619

ABSTRACT

In eye-gaze-based selection, dwell suffers from several issues, e.g., the Midas Touch problem. Here we investigate saccade-based selection techniques as an alternative to dwell. First, we designed a novel user interface (UI) for *Actigaze* and used it with (*goal-crossing*) saccades for confirming the selection of small targets (i.e., < 1.5-2°). We compared it with three other variants of *Actigaze* (with *button press*, *dwell*, and *target reverse crossing*) and two variants of *target magnification* (with *button press* and *dwell*). Magnification-dwell exhibited the most promising performance. For Actigaze, goal-crossing was the fastest option but suffered the most errors. We then evaluated goal-crossing as a primary selection technique for normal-sized targets (≥ 2°) and implemented a novel UI for such interaction. Results revealed that dwell achieved the best performance. Yet, we identified goal-crossing as a good compromise between dwell and button press. Our findings thus identify novel options for gaze-only interaction.

CCS CONCEPTS

• **Human-centered computing → Virtual reality**; *Pointing devices*; *HCI theory, concepts and models.*

KEYWORDS

Eye-Gaze Tracking, Small Targets, Selection Techniques, Virtual Reality, Fitts' Law, Throughput, Activation Methods, Saccade, Target Reverse Crossing

ACM Reference Format:
Aunnoy K Mutasim, Anil Ufuk Batmaz, Moaaz Hudhud Mughrabi, and Wolfgang Stuerzlinger. 2022. Performance Analysis of Saccades for Primary and Confirmatory Target Selection. In *28th ACM Symposium on Virtual Reality Software and Technology (VRST '22), November 29-December 1, 2022, Tsukuba, Japan.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3562939.3565619

1 INTRODUCTION

Many recent Virtual Reality (VR) headsets support eye-tracking [55]. As eye-gaze (called *gaze* here) can move very fast [4], it affords faster pointing for selection tasks [13]. Gaze provides more options for hands-free interaction in VR, increasing a user's sense of presence and embodiment [18]. Also, people with limited muscle control can use gaze as an alternative means of interaction [47]. However, gaze-only interfaces suffers from drawbacks.

One issue is the reliability of selecting/activating a target [71]. The reason is that the same sensory system, the eyes, is used for both perception and control [60]. Without a reliable selection method, gaze is prone to the *Midas Touch* (MT) problem [90], i.e., the unintentional activation of non-targets. *Dwelling*, i.e., fixating ones' gaze for a certain *dwell time*, is most commonly used for gaze-only selection [35, 59, 67, 79]. However, it suffers from the MT problem, especially when the dwell time is low [73], is unnaturally long [83], which limits performance [46], and, it is substantially slower compared to non-gaze-based input methods [19, 27, 37, 72, 77, 96, 97].

The MT problem also poses a major drawback for dwell-based systems, as it prevents free exploration of the user interface (UI) [65]. Thus, researchers have investigated alternatives [27, 72, 75] with a button click typically achieving the best performance [19, 27, 37, 77, 97]. More explicit gaze-only selection techniques, often called *gaze gestures*, have also been studied [12, 29, 39, 43, 49, 66, 78, 82, 89].

Many such gestures involve one or more saccades (e.g., [29, 39, 66]), i.e., ballistic eye movements [73, 79]. A saccade with an amplitude of ≈15-20° takes only 30-120 ms. Saccade latencies range 100-200 ms before the eyes start to move to the next area of interest [24]. Thus, it takes ($SaccadeTime + Latency$ = 30-120 ms + 100-200 ms =) 130-320 ms to complete a single saccade. Also, unlike dwell, using saccades for selection gives users more explicit control over confirming that selection. Thus, saccades are a good potential alternative, even with an expert-level dwell time of 300 ms [59].

Another challenge for gaze-only systems is the limited accuracy and precision of eye-trackers [23]. Interacting with targets less than 1.5-2° with gaze is challenging as the error rate increases substantially [31, 93]. Targets should thus span at least 1.5-2° [79]. Yet, smaller interface elements, < 1.5-2°, provide more options for UI designers, e.g., for small menu items [74], website links [56], overlapping targets [51, 75, 82, 84], or code debugging [81].

Several techniques have been proposed to interact with targets < 1.5-2° (e.g., [32, 56, 75, 81]), many of which are based on 2D target magnification [2, 6, 20, 30, 34, 48, 52, 63, 66, 68, 92]. Yet, selection in immersive systems is more challenging. For instance, since VR headsets are tracked by external sensors, the position and rotation data are subject to jitter. As eye-trackers are embedded in VR headsets, the data collected includes multiple jitter sources, which have detrimental effects on user performance [9–11, 38].

Lutteroth et al. [56] proposed a gaze-only solution, *Actigaze*, to interact with small targets (i.e., < 1.5-2°). Similar to previous work on magnification [2, 6, 20, 48, 52, 68], Actigaze relies on a two-step process, where a selection is followed by a confirmation. The user first dwells near the target – in their case, text-based weblinks. This activates a set of secondary *confirm buttons* (CBs) that share the same color as the weblinks close to the current gaze point. Dwelling on one of the CBs then selects the corresponding target.

Beside dwell, another option for target selection is *goal-crossing* [1], where the cursor crossing an edge of the target selects it. A similar technique, *target reverse crossing* [29], activates a target when the cursor enters and exits through the same edge/arc.

To address some of the most prominent gaps in the literature, we examined the following research questions: *How is user performance affected by saccade-based selection confirmation for small target selection in VR? How does Actigaze compare to the technique of magnifying the targets? As gaze-only alternatives, how well do saccade-based goal-crossing and target reverse crossing perform, compared to button click or dwell?* and *Do these selection techniques differ in performance for primary and confirmatory target selection?*

To answer these research questions, we conducted two user studies. First, we investigated the potential of saccades as the confirmation step selection method. Although saccades can be faster than dwell, saccadic selection by just looking at the target is more susceptible to the MT problem. Yet, previous work [81] argued that, if the selection involved a two-step process, this approach could reduce the MT problem (to some extent), but at the cost of having to take two separate actions to activate a single target. Thus, a two-step process selection mechanism with saccades could offer a good compromise between selection speed and the MT problem.

To evaluate this idea, we use a Fitts' law task [7, 19, 70, 79], with closely packed 0.7°, 0.85°, or 1° targets (see Figure 1). The targets were selected either with one of four variants of Actigaze or two variants of *4x magnification* (4xM). The four Actigaze variants are: activating the CBs using a *button press* (ABP), *dwell* (ADw), *goal-crossing* (AGC), or *target reverse crossing* (ARC). The two variants of 4xM activate confirmation step targets either by a *button press* (4xMBP) or *dwell* (4xMDw), with 4x magnification.

In a second study, we compared saccade-based selection methods, *goal-crossing* and *target reverse crossing*, to *dwell* and *button press* for normal-sized targets (i.e., ≥ 2°), as the primary step selection technique, in an ISO 9241-411 Fitts' law task [3, 40, 71, 72].

Our main contributions are: (1) We present performance measurements for six selection techniques – ABP, ADw, AGC, ARC, 4xMBP, and 4xMDw. Also, we analyze the performance of the corresponding confirmation step activation methods in detail. (2) Similarly, we measure the performance of button press, dwell, goal-crossing, and target reverse crossing as primary step activation methods. (3) We also present two novel UIs. First, a new layout design variation for Actigaze, useful for menus, text entry, and other similar UI items (see Figure 1). Second, a new UI to facilitate saccadic selection as the primary step activation method. For this, we added a larger (otherwise hidden) confirm button (CB) close to each target, which pops up whenever the users' gaze cursor is on the target. Then, making a saccade towards the CB activates the target (see Figure 5).

2 LITERATURE REVIEW

2.1 Fitts' Law and Gaze Tracking

Recently, and matching previous work [87, 95, 97], Schuetz et al. [79] showed that Fitts' law applies to gaze, especially since secondary "corrective" gaze movements are typically required after a main target-directed saccade to point to a target. Fitts' law-based tasks have also been used in the past to study gaze-based selection [79], on 2D screens [17, 41, 79], and in head-mounted displays (HMDs), both in 2D [3, 27, 35] and 3D [72, 76]. The purpose of such studies was to compare different interaction techniques. Qian and Teather [76] compared three types of "gaze" in a VR HMD – eye-gaze, head-gaze, and a combination of the two (eye/head-gaze). The last one corresponds best to real-world gaze behaviours and we thus refer to the eye/head-gaze condition [76] simply as "gaze" [13].

Here, we use the Shannon formulation [57] of Fitts' law [5, 16, 22, 45, 69, 85, 91]. More specifically, we use Kopper et al.'s [45] angular *ID* formula and their method to convert Euclidean distances to angular measures. We assess users' throughput performance also with ISO 9241-411:2012 throughput [40]. A detailed description and the calculation steps for throughput are documented in previous work, e.g, [7, 38, 57, 71].

2.2 Gaze-Based Interaction with Small Targets

The conventional way of interacting with small targets is magnifying the target grid, which involves a two-step process – a magnification followed by a confirmation step. Lankford [52] used dwell to trigger the magnification. Ashmore et al. [2] used a fisheye lens to facilitate zooming. Other inputs for triggering magnification has been explored [6, 44, 48]. Mott and Wobbrock [68] used a Bubble Lens (BL) to enlarge small targets. Using Meyer's model [62], the authors automatically triggered the magnifying lens based on the gaze cursor velocity to create a gaze-only UI. Different types of cursors have also been investigated for small target selection [30]. Recently, Choi et al. [20] compared two gaze-based 2D area cursors — Bubble Cursor (BC) [32] and Bubble Gaze Lens (BGL; a fusion of BC and BL). The authors found that BGL performed best.

Previous work also explored gaze-based target disambiguation [51, 75, 82, 84]. Lutteroth et al. [56] presented a novel gaze-only click alternative, Actigaze, to interact with small targets. Here the user designates a target by first dwelling for 80 ms near any clickable target. This assigns a (limited number of) *confirm button* (CB) colors to all nearby clickables. Selecting the respective CB (at a screen margin) by dwelling on it for 200 ms then selects the desired target. CodeGazer [81] improved this design by placing the CBs in both left and right margins and increased the CB activation time to 300 ms. Since a two-step process is more reliable [81], we decided to experiment with different activation methods for the confirmation step of Actigaze and 4xM, to potentially speed up selection. Also, we build on the idea of a secondary CB [56, 81] and evaluate saccade-based target selection for the primary step.

2.3 Activation Methods for Gaze-Based Systems

Although dwell is popular for activation in gaze systems [21, 35, 67, 79], it suffers from shortcomings [46, 65, 67, 83]. To improve its performance in dwell-based gaze typing, Mott et al. [67] introduced a system that dynamically adjusts the dwell time, depending on the probability that a particular key might be selected. Similarly, Isomoto et al. [41] used Fitts' Law estimates to reduce the dwell time. Researchers also explored alternative activation methods, including button clicks, speech, hand gestures, eye blinks, and electromyography (EMG) [17, 27, 42, 54, 71, 72]. Several studies identified that a button press outperforms other alternatives [19, 27, 37, 77, 97].

Past work also explored *gaze gestures* to activate a target [75, 82], based on one or more saccades [39, 64] or on goal-crossing [1, 26, 65, 66, 73, 89]. One approach involves *target reverse crossing*, where a target is activated when the gaze cursor crosses the same edge/arc twice [29]. Compared to dwell, this requires less time to select a target at the cost of reduced accuracy [29]. Similar techniques have been used in different types of applications [39, 49, 50, 78]. Following this, we also explore saccade-based activation as a primary and confirmation step selection technique.

3 USER STUDY 1

3.1 Conditions and Implementation

3.1.1 Actigaze. We modified Actigaze in several ways in our implementation. Following previous work [6, 30, 44, 48], we use a button press for the first step to activate the 3x3 grid, instead of dwell [56, 81]. Even though this could be replaced by an automatic invocation method [68] to make this a gaze-only UI, we chose the more reliable and faster method with the spacebar [19, 27, 37, 77, 97]. For our 6x6 target grid, where the targets are closely packed together, dwell would not be appropriate for the first step as there are multiple potential overlapping options for the 3x3 grid. Thus, there is an ambiguity as to when to start and stop the dwell timer. Also, cursor jitter would further complicate dwell especially if individual buttons need to be dwelled on to activate the 3x3 grid.

Given that all targets were arranged in a 6x6 grid in our study, the user can be at most one target off when invoking the second step of Actigaze. In other words, the users' gaze cursor needs to be at most one target size distance (i.e. 0.7°, 0.85°, or 1°) away from the target when they hit the spacebar to complete the first step. Otherwise, that particular target would result immediately in a "miss". This means that the user has to be sufficiently accurate in the first step ($targetSize \times 3$) – in essence 2.1°, 2.55°, or 3° for the three target sizes, respectively. If the user hit the spacebar while their gaze was close enough to the target, we assigned CB colors (from the Tableau 20 palette [86]) to the 3x3 sub-grid centered on the button that the user gazed at. If the user selected a corner button, a 2x2 sub-grid at the corner was colored. We also added a yellow "halo" around the desired target in the 3x3 sub-grid to clearly indicate the target.

After this step, users selected the desired CB in one of four ways (see below). The size of the CBs was either 2.8°, 3.4°, or 4° for target sizes 0.7°, 0.85°, or 1° respectively. Placing the circle of CBs closer to the user (see Figure 1) enabled us to use such large(r) CBs. In each case, we provided the corresponding auditory feedback when the correct/incorrect CB was activated. Figure 1 presents a graphical representation of our version of Actigaze.

- *Actigaze-Button Press (ABP)*: A CB is correctly activated if the gaze cursor is on the CB and the spacebar is hit (again).
- *Actigaze-Dwell (ADw)*: To activate a CB, the user has to dwell on it for 300 ms. This also means that if the user dwelled on the wrong CB, this results in a "miss", i.e., the corresponding target in the 6x6 grid would be incorrectly selected (confirmed). We chose 300 ms based on previous work [3, 35, 59, 81].
- *Actigaze-Goal-Crossing (AGC)*: A CB gets activated immediately when the gaze cursor enters it. If the cursor never falls on the CB, e.g., because the sampling frequency is too low, nothing gets selected. Entering the wrong CB (including crossing over it) results in a miss.
- *Actigaze-Target Reverse Crossing (ARC)*: In this version, a CB is activated when the user's gaze enters it from any side but exits only through the inner edge/arc. Although Feng et al.'s [29] original version required the cursor to cross the same edge/arc twice, considering our specific layout, we decided to remove this restriction, as it (unnecessarily) adds more constraints for the user. In essence, the gaze gesture the user needs to perform involves entering a CB and, if they want to confirm it, return back to the 6x6 grid for the next target. However, if the CB entered is not the desired one, they can use a circular motion gaze gesture to correct it, similar to Quikwriting [12]. Entering the *wrong* CB and exiting through its inner edge/arc was counted as a miss.

Compared to Actigaze [56, 81], we arranged the CBs differently in our system and use a CB layout that preserves the relative directions from the main target grid. Thus, the top-left button in the 3x3 sub-grid corresponds to the top-left CB, and so on. Yet, such an arrangement allows for at most eight CBs, leaving no space for a center target CB. Thus, we added two CBs for the center button, at both the top and the bottom, giving the user the choice to either go up or down when they correctly selected the target in the first step. This preserves the directionality of the mapping and also (indirectly) encourages users to gaze as close as possible to the correct target. Also, to avoid unintended confirmations when the cursor just "cuts off a corner" or falls on the edge between two CBs, we added a small gap between them.

3.1.2 4x Magnification. As we implemented the task in a 3D virtual environment, and instead of an explicit magnifying lens, we simply bring a replicated version of the magnified area *closer to the user*, see Figure 2b, which achieves the same magnification effect. To ensure a valid comparison with the Actigaze conditions, we also use a two-step process, with the first being exactly the same as in the Actigaze conditions. Instead of coloring, we magnify a copy of the 3x3 sub-grid for the second confirmation step. Then, the user only needs to point to a target in the magnified grid and confirm the selection. With this design goal-crossing or target reverse crossing cannot be used. Thus, we only evaluate button press and dwell for target activation in the 3x3 sub-grid.

- *4x Magnification with Button Press (4xMBP)*: Pointing at the target in the 3x3 grid and hitting the spacebar (again) activated it.
- *4x Magnification with Dwell (4xMDw)*: Similar to ADw, we chose a 300 ms dwell time. Yet, as the 3x3 grid is magnified towards the user, the gaze cursor will always already hover over one of the buttons in the magnified grid. In our pilots, this caused almost immediate triggering, before users had the time to gaze at the correct target. Thus, we added a delay of 200 ms, i.e., the reaction time [25, 88], so that users have time to perceive the magnified targets and correct their gaze accordingly.

We highlight any button that the user's gaze cursor is in contact with in dark blue. Thus, if the user's gaze cursor was on the target in the first step, when the sub-grid was magnified, that target remained blue in the 3x3 sub-grid. This made it harder to identify the target within the magnified sub-grid. To resolve this, and just like for Actigaze, we added a yellow "halo" around the target in the 3x3 sub-grid. Following previous work [20, 68], we magnified 4x. This also guaranteed that the zoomed targets were the same size as the CBs, i.e., 2.8°, 3.4°, or 4° for target sizes of 0.7°, 0.85°, and 1°, respectively. Thus, 4x zooming ensured that the zoomed targets are easily interactable, as even the smallest zoomed target size is close to previously suggested sizes, i.e., 3° [35, 76, 79]. Finally, just like in Actigaze, the respective notification sounds were played for correct/incorrect selection. See Figure 2 for an illustration of 4xM.

3.2 Hypotheses

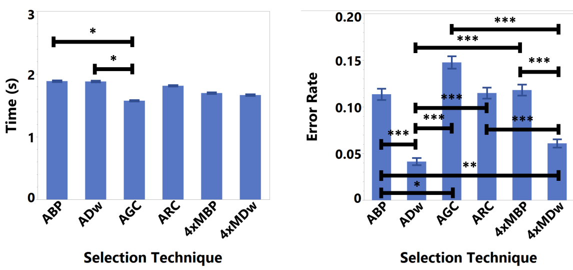
H1.1 Among the four variants of Actigaze, AGC has the highest error rate. The CBs in AGC get activated the instant the cursor is in contact with them. Thus, users can accidentally activate the wrong CB during visual search. **H1.2 AGC is the fastest selection technique for Actigaze.** Goal-crossing does not need additional time like dwell nor the time required to press a button. Also, unlike target reverse crossing, goal-crossing is a one-step mechanism and therefore should save the time required to exit. **H1.3 4xMBP exhibits the best performance among all the conditions.** An advantage of 4xMBP is that the targets are magnified in the direction of the user's gaze. This reduces the gaze travel distance to the target in the 3x3 grid to at most one target size (as the gaze can be at most one target off in the first step). In contrast, the travel distance (for the confirmation step) in Actigaze is larger as the CBs are further away. Also, as a button press is reported to be faster than dwell [19, 27, 37, 77, 97], 4xMBP should outperform 4xMDw.

3.3 Participants, Apparatus, and Procedure

18 participants (4 female) took part in this study. Their mean age was 21.6 ± 3.31 years. All had normal or corrected-to-normal vision.

In this study we used a computer with i7-11700F processor, 32 GB RAM, and a RTX 3070 graphics card, using the Unity game engine. As VR HMD, we used a HTC VIVE Pro Eye HMD with embedded Tobii eye-tracking, with 90 Hz refresh rate, 2880x1600 pixels, and 110° (diagonal) FOV. Its eye-tracker has an accuracy of 0.5-1.1° and transmits data at 120 Hz. Every time participants put the headset (back) on, the eye-tracker was calibrated using Tobii's 5-point calibration method.

Participants started the experiment by filling out a demographic questionnaire. All participants experienced all six conditions, in counterbalanced order. At the start of each of these conditions the experimenter first explained the technique. Participants then put on the VR HMD, where they saw the virtual environment with the target grid presented in front of them (see Figure 1 and 2). After participants had performed some practice trials, they were instructed to do the task as quickly and accurately as possible in the main experiment. They performed the task while sitting on a chair and used the spacebar of a keyboard placed on the desk in front of them. After each condition, they were asked to fill the NASA task load index (NASA-TLX) [36]. At the end of the experiment, subjects filled a short questionnaire where they shared their preferences.



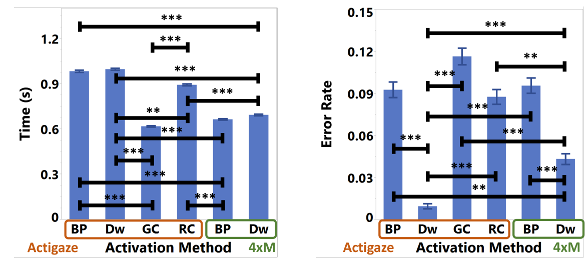
(a) (b)

Figure 3: Time and Error Rate for Selection Techniques.

The center of the 6x6 target grid was placed at the participant's eye level. The experimental task matches previous work [7, 70] where the first target was selected randomly. The next target was then randomly selected with the restriction that it was either 2 or 3 targets to the left, right, top, bottom, or diagonally relative to the previous one, yielding four different target distances (TD) for target sizes 0.7° (TD: 1.40°, 2.10°, 1.98°, and 2.97°), 0.85° (TD: 1.70°, 2.55°, 2.40°, and 3.61°), and 1° (TD: 2°, 3°, 2.83°, and 4.24°). The same target was never selected twice within a round. When there were no more buttons that met the above criteria, the next round of trials was presented. At the beginning of each trial, all the buttons in the 6x6 grid were grey except for the yellow target. If subjects correctly selected a target, it was changed to white, otherwise, black. Each target size was repeated twice for each condition (i.e., two rounds per size). Each condition took about 6-8 minutes. The whole experiment lasted ≈1.5 hours with breaks.

3.4 Experimental Design

We used a within-subjects design with six Selection Techniques (6_{ST} : ABP, ADw, AGC, ARC, 4xMBP, and 4xMDw) and three Target Sizes (3_{TS} : 0.7°, 0.85°, or 1°), all of which were presented in counterbalanced order. For each combination of ST and TS, participants performed two round of trials in the 6x6 target grid. In the analysis, we used the **six STs** as our independent variable. The randomness associated with the next target selection did not guarantee a fixed number of data points for each round of trials. Nonetheless, we collected sufficient data from enough subjects that the data exhibited a uniform distribution for the only independent variable. On average, there were about 30 targets per round of trials yielding about $30 \times 3_{TS} \times 6_{ST} \times 2_{repetitions} = 1080$ data points per subject. As dependent variables, we measured participants' (selection) time (s) – time required to select a target, and error rate (ER) – the ratio of incorrect selections over total number of targets selected. To analyze the influence of the activation methods on the STs, the **four activation methods** (button press, dwell, goal-crossing, and target reverse crossing) was also used as an independent variable.



(a) (b)

Figure 4: Time and Error Rate for Activation Methods.

3.5 Results

Using SPSS 27, we analyzed the data using repeated measures (RM) ANOVA with $\alpha = 0.05$. We considered the data to have a normal distribution when Skewness (S) and Kurtosis (K) values were between ±1.5 [33, 61]. If Mauchly's sphericity test was violated, we applied Huynh-Feldt correction where $\epsilon < 0.75$. For simplicity, we only report statistically significant results here. Post-hoc analyses were conducted with the Bonferroni method. The figures show the means and standard error of means and the significance levels are shown as *** for $p < 0.001$, ** for $p < 0.01$, and * for $p < 0.05$.

3.5.1 Selection Techniques. Results of the RM ANOVA revealed a significant effect of selection techniques on both Time ($F_{5,85} = 3.8$, $p < 0.01$, $\eta^2 = 0.181$) and error rate ($F_{5,85} = 23.4$, $p < 0.001$, $\eta^2 = 0.579$). As shown in Figure 3a, AGC was significantly faster than ABP and ADw. Both dwell conditions, ADw and 4xMDw, exhibited significantly less errors than the other four conditions, see Figure 3b, with no significant differences among the other four, except AGC and ABP, where AGC had more errors than ABP.

3.5.2 Activation Methods. To understand the differences between the activation methods, we analyzed Actigaze's button press, dwell, goal-crossing, and target reverse crossing, along with button press and dwell of 4xM, in execution time and error rate. Note that the data analyzed here is just the data for the second step, i.e., after successful, error-free completion of the first step.

RM ANOVA identified a significant difference between the activation methods for time ($F_{5,85} = 43.3$, $p < 0.001$, $\eta^2 = 0.718$) and error rate ($F_{4.13,70.1} = 26.5$, $p < 0.001$, $\eta^2 = 0.609$). Goal-crossing for Actigaze, and button press and dwell for 4xM, were not significantly different from each other but all three were significantly faster than the other three Actigaze activation methods. Also, dwell for Actigaze was significantly slower than target reverse crossing (see Figure 4a). However, Actigaze-dwell and 4xM-dwell were the most and second-most accurate conditions (see Figure 4b).

3.5.3 Subjective Measures. The average overall task load measured via the NASA-TLX for ABP was (M = 44.1, SD = 20.6), ADw (M = 49.2, SD = 20.9), AGC (M = 59.7, SD = 20.1), ARC (M = 56.1, SD = 21.6), 4xMBP (M = 48.4, SD = 20.9), and for 4xMDw (M = 42.9, SD = 22.1). RM ANOVA identified a significant difference between the selection techniques ($F_{5,85} = 5.71$, $p < 0.001$, $\eta^2 = 0.251$). Post-hoc analysis revealed that only the overall task load of AGC was significantly higher than ABP ($p < 0.001$) and 4xMDw ($p < 0.05$).

In the post-experiment questionnaire, each of 5 out of the 18 participants preferred 4xMBP and 4xMDw, 4 preferred ADw, 3 AGC, 1 ABP, and none ARC. Among 4xM, 10 participants preferred 4xMDw and 8 4xMBP. Reasons why participants preferred the 4xMBP condition were, *"Felt it was the fastest and had the most control over it"*, *"It's the easiest task"*, *"I was able to control the speed which made me faster"*, and *"it was easier than [Actigaze]."* Participants who preferred 4xMDw mentioned *"It was easier to deal when there is no need to press a button"*, *"It demanded less mental and physical effort"*, and *"Because I have time ... I didn't feel pressured."*

Among Actigaze, 7 subjects preferred ADw, 5 AGC, and 3 each for ABP and ARC. Example reasons for the choice of ADw were *"I preferred not clicking on the keyboard"* and *"I could not choose instantly the color [which] was mentally hard."* Similarly, for AGC subjects mentioned *"It was easier and more convenient [that it was] immediately selected [making] it faster"* and *"Less effort required."*

3.6 Study 1 Discussion

We evaluated Actigaze with button press (ABP), dwell (ADw), goal-crossing (AGC), and target reverse crossing (ARC), and 4x magnification (4xM) with button press (4xMBP) and dwell (4xMDw) for selection of targets less than 1.5-2° with gaze in VR HMDs.

While there are many significant differences in the results, we discuss here only the most salient findings. Unless stated differently, all results for individual techniques also hold for the respective activation steps. For Actigaze, no significant differences were observed between ADw and ABP, except that ADw exhibited significantly fewer errors than ABP (see Figure 3). Thus, *ADw seems better than ABP*, contradicting previous work [19, 27, 37, 77, 97]. The likely reason is that these studies used a dwell time much longer than 300 ms. Thus, we acknowledge that our results only hold for UIs with a short dwell time of 300 ms (e.g., [59, 81]). Also, for UIs that require a longer dwell time, one could add an extra (say) 150 ms to our results to arrive at a good estimate of that technique's potential performance.

AGC was significantly faster than ABP and ADw but suffered from more errors than both. Yet, no significant difference in terms of error rate was observed between AGC and ARC (see Figure 3). Thus, we conclude that these results partially support hypothesis **H1.1**, i.e., that AGC has the highest error rate. Given the first step of the two-step process was same for all the conditions, we further investigated the effect of the confirmation step activation methods on the results. As per Figure 4a, goal-crossing was significantly faster than the other three. Thus, our results support hypothesis **H1.2** that AGC is the fastest Actigaze selection technique. Other than dwell, we did not observe any significant differences for goal-crossing in terms of error-rate (see Figure 4b). Although AGC showed higher task load than ABP, AGC was the second-most preferred Actigaze variant by participants. Given these results, and because of the higher error rate of AGC, we conclude that *AGC is the best performing variant of Actigaze in terms of time albeit with the highest error rate.* These results are promising especially because a gaze-only selection technique was able to outperform button press for time, contradicting previous work [19, 27, 37, 77, 96, 97].

Even though the difference in error rate between AGC and ABP will slow down error-free interaction, e.g., for text entry, due to the time needed for error recovery, we speculate that the performance of the two techniques will still be comparable, as AGC was faster than ABP. At a minimum, AGC, i.e., *saccades, seems to offer the best compromise between button press and dwell*, especially for UIs that need a dwell time longer than 300 ms.

ARC exhibited significantly more errors than ADw, but no significant differences were found between ARC and ADw for time. Yet, target reverse crossing was significantly faster than dwell. This result is in line with previous work [29]. No significant differences were observed between ARC and ABP. Also, these techniques were least preferred. Given these results, we conclude that *ARC is comparable to ABP*, offering more gaze-only options.

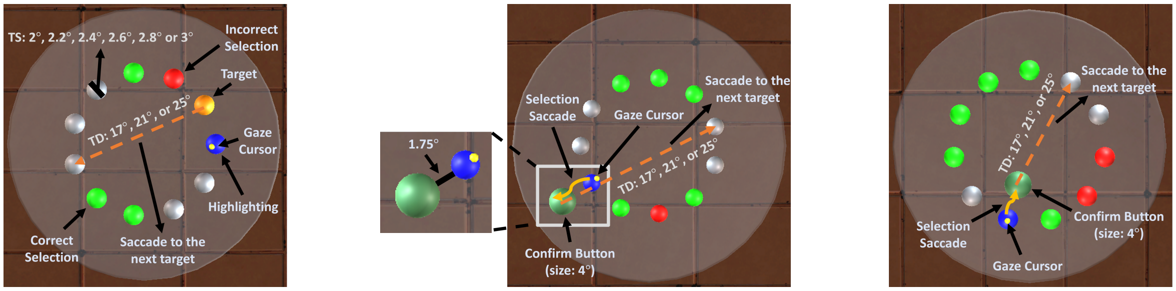
One potential reason why button press was significantly slower than goal-crossing and achieved performance comparable to target reverse crossing might be gaze cursor jitter. We speculate that participants took longer to decide if the cursor will "stay" on the CB at the time they hit the spacebar. In contrast, for target reverse crossing the jitter (sometimes) caused premature exits when hovering on the CB, which then sped up the activation. Based on our observations, this surprised participants, but was not unwelcome.

For the 4xM conditions, 4xMDw exhibited only significantly fewer errors compared to 4xMBP. This means that *4xMDw is the best variant of 4xM*, refuting hypothesis **H1.3**. Given that some application scenarios, such as text entry, could use a short 300 ms dwell with the 4xM interface, we see this as evidence that another gaze-only technique is able to achieve better performance than button press, again contradicting previous work [19, 27, 37, 77, 97].

Comparing 4xMDw with AGC did not reveal any significant differences for time. 4xMDw also exhibited significantly fewer errors than ABP, AGC, and ARC. Dwell of 4xM was also significantly faster than all three Actigaze activation methods except goal-crossing. Although dwell for 4xM exhibited significantly more errors than dwell for Actigaze, no such differences were observed between 4xMDw and ADw (see Figures 3 and 4). Combining the quantitative outcomes with subjects' preferences as well as the significantly lower task load compared to AGC, we conclude that *overall 4xMDw shows the most promising selection performance for small targets.*

No significant differences were observed between 4xMBP and AGC for any of the performance metrics (see Figures 3 and 4). 4xMBP was preferred by 5 participants, i.e., just 2 more participants than AGC. Thus, we also conclude that *both 4xMBP and AGC had comparable performance.*

In line with previous work [19, 27], our results also show that dwell exhibited the least errors for Actigaze and 4xM. However, dwell for 4xM suffered from more errors than for Actigaze (see Figure 3b and 4b). As targets are magnified towards the user in 4xM, the gaze cursor is already in contact with one of the buttons in the magnified grid for the second step which sometimes prematurely activated the button, before the gaze could be moved away.



(a) (b) (c)

Figure 5: The ISO 9241-411 Fitts' law task with (a) button press or dwell, (b) goal-crossing, and (c) target reverse crossing. In (b)-(c) the user looks at the target, which reveals the (hidden) confirm button (CB). A saccade to the CB then selects the target.

4 USER STUDY 2

Goal-crossing and target reverse crossing demonstrated promising performance in Study 1. However, the question remains *whether goal-crossing and target reverse crossing can also achieve similar performance and if they could present alternatives to dwell/button press as a primary step selection technique.* Thus, we decided to investigate them as a primary selection technique with normal-sized targets (i.e., ≥ 2°).

4.1 Conditions and Implementation

- *Button Press (BP)*: If the gaze cursor is on the target when the spacebar is pressed, it is counted as a correct selection.
- *Dwell (Dw)*: This was the same as in the first study – 300 ms dwell time and the conditions for hit and miss.
- *Goal-Crossing (GC)*: To reduce the impact of the MT problem, we implemented a novel UI. The user first gazes at the target. This pops up an otherwise hidden CB "behind" the target relative to the line from the previous CB to the current target. Making a quick saccade to the CB then selects the target (see Figure 5b).
- *Target Reverse Crossing (RC)*: Everything is same as in the GC condition except that the CB appears at a position that required a reverse saccade from the user (see Figure 5c).

For GC and RC, if the gaze cursor left the target, the CBs were (re-)hidden after 50 ms, which gave users 50 ms to hit the CB. As identified in pilot studies, this helped to minimize accidental activations when two CBs appear close to each other. Based on pilots, we also made the size of the CB quite large, i.e., 4°, as saccade paths are not perfectly straight. Otherwise, users had to make repeated saccades to activate the same target. As we noticed premature exits for target reverse crossing in Study 1, we placed the CBs also 1.75° away from the targets, following Patidar et al.'s [73] suggestion. This avoided unintended triggering of CBs solely due to cursor jitter. We chose all these values to be as small as possible based on pilots, while still maintaining an acceptable level of performance.

The target arrangement in Study 1 used too many IDs to enable deeper analysis. Thus, we used the more standardized "circle of targets arrangement", i.e., the ISO 9241-411 task [40], to enhance reliability and replicability. To keep the comparison fair, we modified the target spacing such that the distance from the user's last selection point to the next target was the same across all conditions [57, 58]. For GC and RC, this means that the distance between the (center of the) CB (i.e., the last selection point) and the (center of the) next target is the same as the distance between two target centers for the BP or Dw conditions (see Figure 5).

4.2 Hypotheses

H2.1 GC has the highest throughput among the gaze-only methods. Study 1 showed that, for Actigaze, goal-crossing was fastest. As a single saccade takes only 130-320 ms [24], this option could potentially be faster and therefore, achieve higher throughput than dwell. Finally, Patidar et al. [73] argued that selection would be faster if it required only one or more saccades in the same direction. Thus, GC should achieve higher throughput than RC. **H2.2 The dwell condition has the least error rate.** This is based on the findings of previous work [19, 27] and Study 1.

4.3 Participants, Apparatus, and Procedure

Twelve new participants (6 male, 6 female), aged 23.6 ± 5.11 years, took part in this study. All had normal/corrected-to-normal vision. We used the same apparatus as in Study 1.

Similar to previous work [8, 38, 71], our implementation of the ISO 9241-411 task comprised of 11 targets, placed along the circumference of a circle. The first target in a given "round" of trials was randomly selected. The next targets alternated across the center of the circle, randomly chosen either in a clockwise or anti-clockwise direction (see Figure 5). The experiment comprised of six target sizes between 2° and 3°. Each of these was repeated for three target distances, with a maximum target distance of 25°, based on pilot studies. This minimizes head movements and therefore, makes the experiment less tiring. Each condition took about 7 minutes with the whole experiment lasting about 50 minutes.

All buttons were grey at the beginning of each round, except for the current orange target (see Figure 5a). The target (or any button) was changed to blue whenever the gaze cursor came in contact with it [87]. Upon correct/incorrect selection, the target was changed to green/red, respectively and appropriate notification sounds were played. We kept the rest of the procedure the same as in Study 1.

4.4 Experimental Design

We used a within-subjects design with **four activation methods** (4_{AM} : BP, Dw, GC, and RC), six target sizes (6_{TS} : 2°, 2.2°, 2.4°, 2.6°,

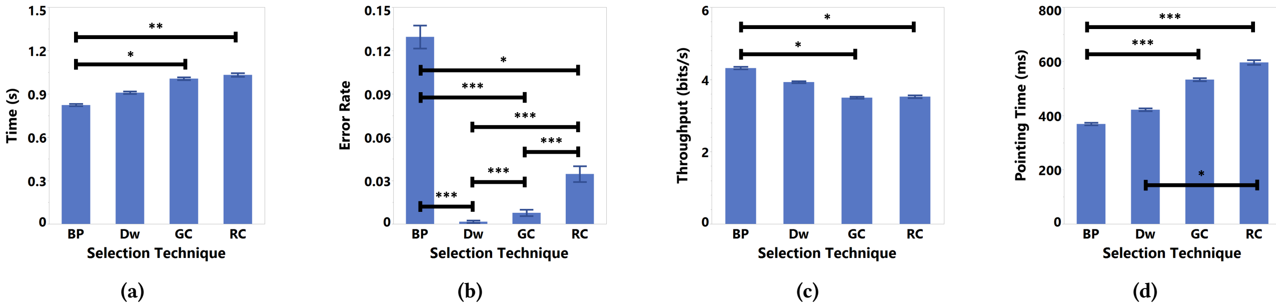


Figure 6: (a) Time, (b) Error Rate, (c) Throughput, and (d) Pointing Time results for Primary Step Selection Techniques.

2.8°, or 3°), and three target distances (3_{TD} : 17°, 21°, or 25°), all of which were presented in counterbalanced order. This gave us a total of **eighteen IDs** ($6_{TS} \times 3_{TD} = 18_{ID}$) between 2.74-3.75. The activation methods and the IDs were independent variables. As dependent variables, we measured participants' (selection) time (s), error rate (ER), effective throughput (THP; bits/s) [40], activation time (ms) – the time taken to activate a target from the last time the gaze cursor has entered the target, total activation time (ms) – the time taken to activate a target after the gaze cursor has reached the target the very first time, and pointing time (ms) - the time taken by the gaze cursor to reach the target. Thus, similar to previous work [29], selection time is the sum of pointing time and the total activation time. There were 11 targets per round yielding $11 \times 6_{TS} \times 3_{TD} \times 4_{AM} = 792$ data points per participant.

4.5 Results

We analyzed the data using RM ANOVA with the same parameters as in Study 1. For dependent variables that did not have a normal/log-normal distribution, data was transformed using Aligned Rank Transform (ART) [94] before RM ANOVA.

4.5.1 Time, Error Rate, and Throughput Analysis. Results of the RM ANOVAs are presented in Table 1. As shown in Figures 6a and 6c, users were significantly faster and had higher throughput in the BP condition compared to GC and RC. Users made the least errors in the Dw condition, followed by GC, RC, and BP, with all significantly different from each other (see Figure 6b).

4.5.2 Activation and Pointing Time Analysis. Average activation time for BP was 297.8 ms, Dw 305.2 ms, GC 268.7 ms, and RC 262.4 ms. The total activation time for BP was 454.0 ms, Dw 487.1 ms, GC 473.5 ms, and RC 435.7 ms. RM ANOVA identified only a significant difference between the selection techniques for pointing time ($F_{1.73,15.6} = 18.9$, $p < 0.001$, $\eta^2 = 0.678$) and ID ($F_{17,153} = 4.68$, $p < 0.001$, $\eta^2 = 0.342$). Post-hoc analysis revealed that BP (369.2 ms) was significantly faster than GC (533.1 ms) and RC (596.5 ms). Dw (422.0 ms) exhibited a faster pointing time than RC (see Figure 6d).

Table 1: Selection Techniques RM ANOVA results.

	Selection Techniques	ID
Time	$F_{3,27} = 6.89$, $p < 0.001$, $\eta^2 = 0.434$	$F_{17,153} = 3.01$, $p < 0.001$, $\eta^2 = 0.251$
ER	$F_{2.21,19.9} = 73.8$, $p < 0.001$, $\eta^2 = 0.891$	$F_{17,153} = 7.71$, $p < 0.001$, $\eta^2 = 0.461$
THP	$F_{2.40,21.6} = 6.13$, $p < 0.01$, $\eta^2 = 0.405$	$F_{17,153} = 2.97$, $p < 0.001$, $\eta^2 = 0.248$

4.5.3 Subjective Measures. The average overall task load measured via the NASA-TLX for BP was ($M = 29.6$, $SD = 19.9$), Dw ($M = 28.4$, $SD = 13.3$), GC ($M = 50.7$, $SD = 21.2$), and RC ($M = 46.4$, $SD = 19.1$). RM ANOVA identified a significant difference between the selection techniques ($F_{3,33} = 12.7$, $p < 0.001$, $\eta^2 = 0.536$). Post-hoc analysis revealed that the overall task load of GC and RC was significantly higher than BP (both $p < 0.05$) and Dw (both $p < 0.01$).

In the post-experiment questionnaire, 6 subjects preferred Dw, 3 BP, 2 RC, and 1 GC. Example reasons given by subjects for their preference of Dw was, *"better performance and easy"*, *"Least amount of effort needed. No frustration,"* and *"The mental activity needed ... was very low."* Subjects who chose BP mentioned *"I had a chance to confirm my choices"*, and *"Felt the easiest and the fastest."* For RC one subject mentioned *"It was the easiest to select since my gaze would naturally go back to the center."* Finally, for GC, *"I enjoyed it more."*

4.6 Study 2 Discussion

Here, we investigated four primary step selection techniques – button press (BP), dwell (Dw), goal-crossing (GC) and target reverse crossing (RC). Results showed that, compared to GC and RC, BP had significantly better performance for time, throughput, and task load. No such difference was observed between BP and Dw (see Figure 6). Further analysis revealed that the reason why BP is superior to GC and RC is that subjects were able to point much faster in BP (see Figure 6d). However, users made the most errors with BP (see Figure 6b). The reason for the high error rate might be that users pressed the button before their gaze reached the target or left the target before they finished hitting the button (as gaze moves very fast). Still, BP's high error rate is in line with previous work [19, 27].

Just like Study 1 and previous work [19, 27], Dw again exhibited the least errors in Study 2, thus supporting hypothesis **H2.2**. Moreover, Dw was not significantly different than BP in any other performance metric. It was also highly preferred by participants with significantly less task load compared to GC and RC. Thus, we conclude that *Dw is the best primary step selection technique* which matches Hansen et al.'s results [35]. Yet, our results contradict the findings of other previous work [19, 27, 37, 77, 97], which reported button press to be superior to dwell. Again, the likely reason is that these studies used a much higher dwell time than 300 ms [35].

The reason behind the low error rate for dwell in both our studies is due to how this technique works. Making a mistake in the dwell condition requires deliberately dwelling on a non-target for 300 ms. In contrast, with button press a miss occurs when the gaze cursor is not on the target (for whatever reason) and the spacebar is hit. This means that it is effectively easier to make a mistake with a button press than dwell. Similarly, for GC, to make a mistake the user needs to hover on the wrong target and then select its CB, which is again harder than BP. Although GC exhibited significantly more errors than Dw, GC's error rate is still very low (< 1%). Except for the task load, GC also did not show any significant difference in the other metrics compared to Dw. Thus, we conclude that *GC is the best compromise between BP and Dw*. Also, since GC did not achieve the highest throughput, the results refute **H2.1**.

Unlike dwell, using saccades to CBs as an explicit selection mechanism enables free visual exploration of the content, e.g., a keyboard layout [65]. Yet, according to our results, saccadic selection is more fatiguing than dwell. Nonetheless, both options have their own set of advantages and thus, we believe that *the choice between dwell and/or saccade depends on the type and the objective of the application*, e.g., when gaze-only is the only viable option.

Compared to GC, a significantly different error rate was observed for RC, but it was still low (see Figure 6b). The potential reason for this is that in RC, the targets were placed further away from the center of the target arrangement to maintain a consistent target distance for Fitts' law. Thus, they were placed further into the periphery where previous work [13, 23, 80] identified lower eye-tracking accuracy. This might also explain why RC was significantly slower than Dw in pointing time (see Figure 6d). Although this slower pointing time did not result in a significant difference for selection time and throughput between Dw and RC, RC was not significantly faster than Dw, contradicting previous work [29] and Study 1. We believe that the smaller dwell time, targets being placed further into the periphery, as well as premature exits of target reverse crossing in Study 1 all contributed to this outcome.

There was no significant difference between GC and RC in selection time, throughput, activation, total activation, and pointing time (see Figure 6). These results refute the previously presented argument that selection would be faster if a technique required only one saccade or more in the same direction [73]. We also found no significant differences in activation and total activation time for the four primary step selection techniques. In other words, **GC and RC as activation methods are comparable to 300 ms dwell, and, more importantly, a button press**, where the latter was previously identified as the fastest option, e.g., [19, 27, 37, 77, 97]. However, we also found that BP was the fastest condition relative to GC and RC in terms of selection time. The difference in speed is thus best explained by the difference in pointing time (see Figure 6d).

Other than the issue of targets being placed further into the periphery for RC (to keep the comparison fair), we speculate that both GC and RC had higher pointing time because there are two saccade latencies involved – one after the previous CB was selected and the other when the target was reached (and the CB is revealed). It could also be that – before deciding to hit the CB with the second saccade – additional cognitive decision time is required due to cursor jitter [38]. Also, visually verifying that the CB appeared when the target is reached might add more latency, too. Further studies are required to verify this.

Our results apply to UIs like Iwrite [89] where keys are selected by saccades in any direction. For UIs with more densely arranged objects, like keyboards, a frame border [89] can be used as a big CB. Separate CBs for each key instead of the two contexts could improve L2's design [66], also saving screen space. Nonetheless, we acknowledge that saccadic selection inherently limits the design of a UI and is thus generally not as suitable as dwell/buttons.

5 GENERAL DISCUSSION

In this paper, we evaluated the performance of saccades for primary and confirmatory target selection. Study 1 identified that 4xM-dwell (4xMDw) achieved the most promising performance for small target selection. Among the Actigaze variants, goal-crossing (AGC) had the best performance for time but suffered the highest error rate. Results from Study 2 revealed that dwell (Dw; 300 ms) exhibited the best performance as the primary step selection technique for normal-sized targets. A saccadic selection technique, goal-crossing (GC), exhibited the best compromise between Dw and button press (BP). Thus, goal-crossing exhibited promising performance both as a primary and confirmatory activation method.

For both studies significant differences exhibited large effect sizes ($\eta^2 > 0.14$) with a number of subjects consistent with much other work [14]. This indicates that our results are likely to be replicable. In the future, it would be interesting to replicate our studies in a 2D system. Such a setup would generalize our results to conventional user interfaces, constitute a more affordable solution, and would also allow researchers to increase the size of the participant pool to further verify our results.

Based on our results, we arrive at the following insights and potential future work for the use of saccades as primary and confirmatory selection techniques. These are informed not only by the outcomes but also on our observations during the studies:

As discussed above, the slower pointing time of GC and RC in Study 2 is likely explained by multiple latencies. We speculate that revealing the confirm button (CB) earlier in the process might speed up selection. Similar to previous work [68], one potential solution would be to **use Meyer's model [62] to reveal the CBs of the target and its immediate neighbors as soon as the gaze cursor is close to the target**. This also means that the CBs would need to stay visible even if the cursor just "falls off" the target due to cursor jitter. However, further studies are required to verify whether this modification would actually yield a significant improvement.

A lot of the mistakes for Actigaze-goal-crossing in Study 1 resulted from a quick visual search or just "cutting off a corner" of a neighboring CB. We thus suggest to **add a small dwell time of 50/100 ms to Actigaze-goal-crossing**, i.e., essentially a version of Actigaze-dwell with a much shorter dwell time, before activating the target. This might significantly reduce the error rate, hopefully without substantially affecting the time results.

That different dwell times are appropriate for different tasks/UIs, e.g., text entry or selecting a target image, key, or icon, is well-documented in the literature, e.g., [19, 21, 27, 37, 67, 77, 97], with shorter dwell times for sparsely arranged objects or longer ones in a dense environment. Yet, all our subjects were novices, i.e., they never trained to reduce their dwell time. Given our results, and in contrast to previous work that claims 300 ms to be only appropriate for experts [59], we thus believe that **300 ms dwell time can be appropriate for novices and expert-level dwell time could be lower for tasks/UIs such as those we evaluated**, which should be investigated in the future.

In both our studies we did not apply a noise filter to the eye-tracking data as we wanted to investigate how well state-of-the-art (VR) eye-trackers work without such filters, especially for small targets. Still, applying a noise filter [15, 28, 53] might noticeably reduce the error rate. Further studies are required to verify this.

6 CONCLUSION

In this work, we investigated whether saccadic selection can serve as an alternate to dwell for primary and confirmatory target selection. We first compared six gaze-based selection techniques for targets smaller than 1.5-2°. Results revealed that 4x magnification with dwell achieved the most promising performance. Actigaze with goal-crossing achieved the best performance for time albeit with the most errors. We also identified in a second study that saccade-based goal-crossing as a primary selection technique is a good compromise between dwell and button press for normal-sized targets. The results of our work enable practitioners, developers, and researchers to make more informed decisions for target selection in gaze-based UIs.

REFERENCES

[1] Johnny Accot and Shumin Zhai. 2002. More than Dotting the i's — Foundations for Crossing-Based Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Minneapolis, Minnesota, USA) *(CHI '02)*. Association for Computing Machinery, New York, NY, USA, 73–80. https://doi.org/10.1145/503376.503390

[2] 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.

[3] Per Bækgaard, John Paulin Hansen, Katsumi Minakata, and I. Scott MacKenzie. 2019. A Fitts' Law Study of Pupil Dilations in a Head-Mounted Display. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research and Applications* (Denver, Colorado) *(ETRA '19)*. Association for Computing Machinery, New York, NY, USA, Article 32, 5 pages. https://doi.org/10.1145/3314111.3319831

[4] A. Terry Bahill, Michael R. Clark, and Lawrence Stark. 1975. The main sequence, a tool for studying human eye movements. *Mathematical Biosciences* 24, 3 (1975), 191 – 204. https://doi.org/10.1016/0025-5564(75)90075-9

[5] Mayra Donaji Barrera Machuca and Wolfgang Stuerzlinger. 2019. The Effect of Stereo Display Deficiencies on Virtual Hand Pointing,. In *2019 CHI Conference on Human Factors in Computing Systems*. ACM, 14 Pages.

[6] Richard Bates and Howell Istance. 2002. Zooming Interfaces! Enhancing the Performance of Eye Controlled Pointing Devices. In *Proceedings of the Fifth International ACM Conference on Assistive Technologies* (Edinburgh, Scotland) *(Assets '02)*. Association for Computing Machinery, New York, NY, USA, 119–126. https://doi.org/10.1145/638249.638272

[7] Anil Ufuk Batmaz, Aunnoy K Mutasim, Morteza Malekmakan, Elham Sadr, and Wolfgang Stuerzlinger. 2020. Touch the Wall: Comparison of Virtual and Augmented Reality with Conventional 2D Screen Eye-Hand Coordination Training Systems. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 184–193. https://doi.org/10.1109/VR46266.2020.00037

[8] Anil Ufuk Batmaz, Aunnoy K Mutasim, and Wolfgang Stuerzlinger. 2020. Precision vs. Power Grip: A Comparison of Pen Grip Styles for Selection in Virtual Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 23–28.

[9] Anil Ufuk Batmaz, Mohammad Rajabi Seraji, Johanna Kneifel, and Wolfgang Stuerzlinger. 2020. No Jitter Please: Effects of Rotational and Positional Jitter on 3D Mid-Air Interaction. In *Future Technologies Conference (FTC '20, Vol. AISC 1289)*. 16. https://doi.org/10.1007/978-3-030-63089-8_52

[10] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2019. The Effect of Rotational Jitter on 3D Pointing Tasks. Extended Abstract. , 6 pages. https://doi.org/10.1145/3290607.3312752

[11] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2019. Effects of 3D Rotational Jitter and Selection Methods on 3D Pointing Tasks. , 1687-1692 pages. https://doi.org/10.1109/VR.2019.8798038

[12] Nikolaus Bee and Elisabeth André. 2008. Writing with Your Eye: A Dwell Time Free Writing System Adapted to the Nature of Human Eye Gaze. In *Perception in Multimodal Dialogue Systems*, Elisabeth André, Laila Dybkjær, Wolfgang Minker, Heiko Neumann, Roberto Pieraccini, and Michael Weber (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 111–122.

[13] Jonas Blattgerste, Patrick Renner, and Thies Pfeiffer. 2018. Advantages of Eye-Gaze over Head-Gaze-Based Selection in Virtual and Augmented Reality under Varying Field of Views. In *Proceedings of the Workshop on Communication by Gaze Interaction* (Warsaw, Poland) *(COGAIN '18)*. Association for Computing Machinery, New York, NY, USA, Article 1, 9 pages. https://doi.org/10.1145/3206343.3206349

[14] Kelly Caine. 2016. Local Standards for Sample Size at CHI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 981–992. https://doi.org/10.1145/2858036.2858498

[15] Géry Casiez, Nicolas Roussel, and Daniel Vogel. 2012. 1€ filter: a simple speed-based low-pass filter for noisy input in interactive systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2527–2530.

[16] Yeonjoo Cha and Rohae Myung. 2013. Extended Fitts' law for 3D pointing tasks using 3D target arrangements. *International Journal of Industrial Ergonomics* 43, 4 (2013), 350 – 355. https://doi.org/10.1016/j.ergon.2013.05.005

[17] Ishan Chatterjee, Robert Xiao, and Chris Harrison. 2015. Gaze+Gesture: Expressive, Precise and Targeted Free-Space Interactions. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (Seattle, Washington, USA) *(ICMI '15)*. Association for Computing Machinery, New York, NY, USA, 131–138. https://doi.org/10.1145/2818346.2820752

[18] Pascal Chiu, Kazuki Takashima, Kazuyuki Fujita, and Yoshifumi Kitamura. 2019. Pursuit Sensing: Extending Hand Tracking Space in Mobile VR Applications. In *Symposium on Spatial User Interaction* (New Orleans, LA, USA) *(SUI '19)*. Association for Computing Machinery, New York, NY, USA, Article 1, 5 pages. https://doi.org/10.1145/3357251.3357578

[19] Mungyeong Choe, Yeongcheol Choi, Jaehyun Park, and Hyun K. Kim. 2019. Comparison of Gaze Cursor Input Methods for Virtual Reality Devices. *International Journal of Human–Computer Interaction* 35, 7 (2019), 620–629. https://doi.org/10.1080/10447318.2018.1484054 arXiv:https://doi.org/10.1080/10447318.2018.1484054

[20] Myungguen Choi, Daisuke Sakamoto, and Tetsuo Ono. 2020. Bubble Gaze Cursor + Bubble Gaze Lens: Applying Area Cursor Technique to Eye-Gaze Interface. In *ACM Symposium on Eye Tracking Research and Applications* (Stuttgart, Germany) *(ETRA '20 Full Papers)*. Association for Computing Machinery, New York, NY, USA, Article 11, 10 pages. https://doi.org/10.1145/3379155.3391322

[21] Myungguen Choi, Daisuke Sakamoto, and Tetsuo Ono. 2022. Kuiper Belt: Utilizing the "Out-of-Natural Angle" Region in the Eye-Gaze Interaction for Virtual Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 357, 17 pages. https://doi.org/10.1145/3491102.3517725

[22] Logan D. Clark, Aakash B. Bhagat, and Sara L. Riggs. 2020. Extending Fitts' law in three-dimensional virtual environments with current low-cost virtual reality technology. *International Journal of Human-Computer Studies* 139 (2020), 102413. https://doi.org/10.1016/j.ijhcs.2020.102413

[23] Viviane Clay, Peter König, and Sabine Koenig. 2019. Eye Tracking in Virtual Reality. *Journal of Eye Movement Research* 12 (04 2019). https://doi.org/10.16910/jemr.12.1.3

[24] Andrew T Duchowski. 2017. *Eye tracking methodology: Theory and practice*. Springer.

[25] Michael D'Zmura. 1991. Color in visual search. *Vision Research* 31, 6 (1991), 951–966. https://doi.org/10.1016/0042-6989(91)90203-H

[26] Carlos Elmadjian and Carlos H Morimoto. 2021. *GazeBar: Exploiting the Midas Touch in Gaze Interaction*. Association for Computing Machinery, New York, NY, USA. https://doi-org.proxy.lib.sfu.ca/10.1145/3411763.3451703

[27] 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. https://doi.org/10.1016/j.ijhcs.2020.102414

[28] 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

[29] Wenxin Feng, Ming Chen, and Margrit Betke. 2014. Target Reverse Crossing: A Selection Method for Camera-Based Mouse-Replacement Systems. In *Proceedings of the 7th International Conference on PErvasive Technologies Related to Assistive Environments* (Rhodes, Greece) *(PETRA '14)*. Association for Computing Machinery, New York, NY, USA, Article 39, 4 pages. https://doi.org/10.1145/2674396.2674443

[30] Leah Findlater, Alex Jansen, Kristen Shinohara, Morgan Dixon, Peter Kamb, Joshua Rakita, and Jacob O. Wobbrock. 2010. Enhanced Area Cursors: Reducing Fine Pointing Demands for People with Motor Impairments. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology* (New York, New York, USA) *(UIST '10)*. Association for Computing Machinery, New York, NY, USA, 153–162. https://doi.org/10.1145/1866029.1866055

[31] Sven-Thomas Graupner, Michael Heubner, Sebastian Pannasch, and Boris M. Velichkovsky. 2008. Evaluating Requirements for Gaze-Based Interaction in a See-through Head Mounted Display. In *Proceedings of the 2008 Symposium on Eye*

Tracking Research and Applications (Savannah, Georgia) *(ETRA '08)*. Association for Computing Machinery, New York, NY, USA, 91–94. https://doi.org/10.1145/1344471.1344495

[32] Tovi Grossman and Ravin Balakrishnan. 2005. The Bubble Cursor: Enhancing Target Acquisition by Dynamic Resizing of the Cursor's Activation Area. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, Oregon, USA) *(CHI '05)*. Association for Computing Machinery, New York, NY, USA, 281–290. https://doi.org/10.1145/1054972.1055012

[33] Joseph F. Hair Jr, William C. Black, Barry J. Babin, and Rolph E. Anderson. 2014. Multivariate data analysis.

[34] Dan Witzner Hansen, Henrik H. T. Skovsgaard, John Paulin Hansen, and Emilie Møllenbach. 2008. Noise Tolerant Selection by Gaze-Controlled Pan and Zoom in 3D. In *Proceedings of the 2008 Symposium on Eye Tracking Research and Applications* (Savannah, Georgia) *(ETRA '08)*. Association for Computing Machinery, New York, NY, USA, 205–212. https://doi.org/10.1145/1344471.1344521

[35] John Paulin Hansen, Vijay Rajanna, I. Scott MacKenzie, and Per Bækgaard. 2018. A Fitts' Law Study of Click and Dwell Interaction by Gaze, Head and Mouse with a Head-Mounted Display. In *Proceedings of the Workshop on Communication by Gaze Interaction* (Warsaw, Poland) *(COGAIN '18)*. Association for Computing Machinery, New York, NY, USA, Article 7, 5 pages. https://doi.org/10.1145/3206343.3206344

[36] Sandra G Hart. 1986. NASA Task load Index (TLX). Volume 1.0; Paper and pencil package. (1986).

[37] Mehedi Hassan, John Magee, and I. Scott MacKenzie. 2019. A Fitts' law evaluation of hands-free and hands-on input on a laptop computer. In *International Conference on Human-Computer Interaction*. Springer, 234–249.

[38] Moaaz Hudhud Mughrabi, Aunnoy K Mutasim, Wolfgang Stuerzlinger, and Anil Ufuk Batmaz. 2022. My Eyes Hurt: Effects of Jitter in 3D Gaze Tracking. In *Conference on Virtual Reality and 3D User Interfaces, Abstracts and Workshops, Workshop on Novel Input Devices and Interaction Techniques at VR (NIDIT '22)*.

[39] Aulikki Hyrskykari, Howell Istance, and Stephen Vickers. 2012. Gaze Gestures or Dwell-Based Interaction?. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (Santa Barbara, California) *(ETRA '12)*. Association for Computing Machinery, New York, NY, USA, 229–232. https://doi.org/10.1145/2168556.2168602

[40] ISO 9241-411:2012. 2012. Ergonomics of human-system interaction – Part 411: Evaluation methods for the design of physical input devices. ISO. https://www.iso.org/standard/54106.html

[41] Toshiya Isomoto, Toshiyuki Ando, Buntarou Shizuki, and Shin Takahashi. 2018. Dwell Time Reduction Technique Using Fitts' Law for Gaze-Based Target Acquisition. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research and Applications* (Warsaw, Poland) *(ETRA '18)*. Association for Computing Machinery, New York, NY, USA, Article 26, 7 pages. https://doi.org/10.1145/3204493.3204532

[42] Janis G Jimenez and Jürgen P Schulze. 2018. Continuous-Motion Text Input in Virtual Reality. *Electronic Imaging* 2018, 3 (2018), 450–1.

[43] Florian Jungwirth, Michael Haslgrübler, and Alois Ferscha. 2018. Contour-Guided Gaze Gestures: Using Object Contours as Visual Guidance for Triggering Interactions. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications* (Warsaw, Poland) *(ETRA '18)*. Association for Computing Machinery, New York, NY, USA, Article 28, 10 pages. https://doi.org/10.1145/3204493.3204530

[44] Konstantin Klamka, Andreas Siegel, Stefan Vogt, Fabian Göbel, Sophie Stellmach, and Raimund Dachselt. 2015. Look & Pedal: Hands-Free Navigation in Zoomable Information Spaces through Gaze-Supported Foot Input. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (Seattle, Washington, USA) *(ICMI '15)*. Association for Computing Machinery, New York, NY, USA, 123–130. https://doi.org/10.1145/2818346.2820751

[45] Regis Kopper, Doug A Bowman, Mara G Silva, and Ryan P McMahan. 2010. A human motor behavior model for distal pointing tasks. *International journal of human-computer studies* 68, 10 (2010), 603–615.

[46] Per Ola Kristensson and Keith Vertanen. 2012. The Potential of Dwell-Free Eye-Typing for Fast Assistive Gaze Communication. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (Santa Barbara, California) *(ETRA '12)*. Association for Computing Machinery, New York, NY, USA, 241–244. https://doi.org/10.1145/2168556.2168605

[47] Chandan Kumar, Raphael Menges, Daniel Müller, and Steffen Staab. 2017. Chromium Based Framework to Include Gaze Interaction in Web Browser. In *Proceedings of the 26th International Conference on World Wide Web Companion* (Perth, Australia) *(WWW '17 Companion)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 219–223. https://doi.org/10.1145/3041021.3054730

[48] 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

[49] Andrew Kurauchi, Wenxin Feng, Ajjen Joshi, Carlos Morimoto, and Margrit Betke. 2016. EyeSwipe: Dwell-Free Text Entry Using Gaze Paths. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1952–1956. https://doi.org/10.1145/2858036.2858335

[50] Andrew Kurauchi, Wenxin Feng, Ajjen Joshi, Carlos H. Morimoto, and Margrit Betke. 2020. Swipe&Switch: Text Entry Using Gaze Paths and Context Switching. In *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 84–86. https://doi.org/10.1145/3379350.3416193

[51] Mikko Kytö, Barrett Ens, Thammathip Piumsomboon, Gun A. Lee, and Mark Billinghurst. 2018. *Pinpointing: Precise Head- and Eye-Based Target Selection for Augmented Reality*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi-org.proxy.lib.sfu.ca/10.1145/3173574.3173655

[52] Chris Lankford. 2000. Effective Eye-Gaze Input into Windows. In *Proceedings of the 2000 Symposium on Eye Tracking Research and 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

[53] Zhi Li, Maozheng Zhao, Yifan Wang, Sina Rashidian, Furqan Baig, Rui Liu, Wanyu Liu, Michel Beaudouin-Lafon, Brooke Ellison, Fusheng Wang, et al. 2021. Bayesgaze: A bayesian approach to eye-gaze based target selection. In *Proceedings. Graphics Interface (Conference)*, Vol. 2021. NIH Public Access, 231.

[54] Xueshi Lu, Difeng Yu, Hai-Ning Liang, Wenge Xu, Yuzheng Chen, Xiang Li, and Khalad Hasan. 2020. Exploration of Hands-free Text Entry Techniques For Virtual Reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 344–349. https://doi.org/10.1109/ISMAR50242.2020.00061

[55] Francisco Lopez Luro and Veronica Sundstedt. 2019. A Comparative Study of Eye Tracking and Hand Controller for Aiming Tasks in Virtual Reality. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research and Applications* (Denver, Colorado) *(ETRA '19)*. Association for Computing Machinery, New York, NY, USA, Article 68, 9 pages. https://doi.org/10.1145/3317956.3318153

[56] 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

[57] I. Scott MacKenzie. 1992. Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction* 7, 1 (1992), 91–139.

[58] I. Scott MacKenzie. 2015. Fitts' Throughput and the Remarkable Case of Touch-Based Target Selection. In *Human-Computer Interaction: Interaction Technologies*, Masaaki Kurosu (Ed.). Springer International Publishing, Cham, 238–249.

[59] Päivi Majaranta, Ulla-Kaija Ahola, and Oleg Špakov. 2009. Fast Gaze Typing with an Adjustable Dwell Time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 357–360. https://doi.org/10.1145/1518701.1518758

[60] Päivi Majaranta and Andreas Bulling. 2014. *Eye Tracking and Eye-Based Human–Computer Interaction*. Springer London, London, 39–65. https://doi.org/10.1007/978-1-4471-6392-3_3

[61] Paul Mallery and Darren George. 2003. SPSS for Windows step by step: a simple guide and reference. *Allyn, Bacon, Boston,* (2003).

[62] David Meyer, J. Keith-Smith, Sylvan Kornblum, Richard Abrams, and Charles Wright. 1990. Speed-Accuracy tradeoffs in aimed movements: Toward a theory of rapid voluntary action. *Attention and Performance XIII* (01 1990).

[63] 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

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

[65] Carlos H. Morimoto and Arnon Amir. 2010. Context Switching for Fast Key Selection in Text Entry Applications. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (Austin, Texas) *(ETRA '10)*. Association for Computing Machinery, New York, NY, USA, 271–274. https://doi.org/10.1145/1743666.1743730

[66] Carlos H. Morimoto, Jose A. T. Leyva, and Antonio Diaz-Tula. 2018. Context Switching Eye Typing Using Dynamic Expanding Targets. In *Proceedings of the Workshop on Communication by Gaze Interaction* (Warsaw, Poland) *(COGAIN '18)*. Association for Computing Machinery, New York, NY, USA, Article 6, 9 pages. https://doi.org/10.1145/3206343.3206347

[67] Martez E. Mott, Shane Williams, Jacob O. Wobbrock, and Meredith Ringel Morris. 2017. Improving Dwell-Based Gaze Typing with Dynamic, Cascading Dwell Times. 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, 2558–2570. https://doi.org/10.1145/3025453.3025517

[68] Martez E. Mott and Jacob O. Wobbrock. 2014. Beating the Bubble: Using Kinematic Triggering in the Bubble Lens for Acquiring Small, Dense Targets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 733–742. https://doi.org/10.1145/2556288.2557410

[69] Atsuo Murata and Hirokazu Iwase. 2001. Extending Fitts' law to a three-dimensional pointing task. *Human Movement Science* 20, 6 (2001), 791 – 805. https://doi.org/10.1016/S0167-9457(01)00058-6

[70] Aunnoy K Mutasim, Anil Ufuk Batmaz, and Wolfgang Stuerzlinger. 2020. Gaze Tracking for Eye-Hand Coordination Training Systems in Virtual Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3334480.3382924

[71] Aunnoy K Mutasim, Anil Ufuk Batmaz, and Wolfgang Stuerzlinger. 2021. Pinch, Click, or Dwell: Comparing Different Selection Techniques for Eye-Gaze-Based Pointing in Virtual Reality. In *ACM Symposium on Eye Tracking Research and Applications* (Virtual Event, Germany) *(ETRA '21 Short Papers)*. Association for Computing Machinery, New York, NY, USA, Article 15, 7 pages. https://doi.org/10.1145/3448018.3457998

[72] Yun Suen Pai, Tilman Dingler, and Kai Kunze. 2019. Assessing hands-free interactions for VR using eye gaze and electromyography. *Virtual Reality* 23, 2 (2019), 119–131.

[73] Pawan Patidar, Himanshu Raghuvanshi, and Sayan Sarcar. 2014. Quickpie: An Interface for Fast and Accurate Eye Gazed based Text Entry. arXiv:1407.7313 [cs.HC]

[74] 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

[75] Thammathip Piumsomboon, Gun Lee, Robert W Lindeman, and Mark Billinghurst. 2017. Exploring natural eye-gaze-based interaction for immersive virtual reality. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. 36–39.

[76] Yuan Yuan Qian and Robert J. Teather. 2017. The Eyes Don't Have It: An Empirical Comparison of Head-Based and Eye-Based Selection in Virtual Reality. In *Proceedings of the 5th Symposium on Spatial User Interaction* (Brighton, United Kingdom) *(SUI '17)*. Association for Computing Machinery, New York, NY, USA, 91–98. https://doi.org/10.1145/3131277.3132182

[77] Vijay Rajanna and John Paulin Hansen. 2018. Gaze Typing in Virtual Reality: Impact of Keyboard Design, Selection Method, and Motion. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research and Applications* (Warsaw, Poland) *(ETRA '18)*. Association for Computing Machinery, New York, NY, USA, Article 15, 10 pages. https://doi.org/10.1145/3204493.3204541

[78] Sayan Sarcar, Prateek Panwar, and Tuhin Chakraborty. 2013. EyeK: An Efficient Dwell-Free Eye Gaze-Based Text Entry System. In *Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction* (Bangalore, India) *(APCHI '13)*. Association for Computing Machinery, New York, NY, USA, 215–220. https://doi.org/10.1145/2525194.2525288

[79] Immo Schuetz, T. Scott Murdison, Kevin J. MacKenzie, and Marina Zannoli. 2019. An Explanation of Fitts' Law-like Performance in Gaze-Based Selection Tasks Using a Psychophysics Approach. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300765

[80] Immo Schuetz, T. Scott Murdison, and Marina Zannoli. 2020. A Psychophysics-Inspired Model of Gaze Selection Performance. In *ACM Symposium on Eye Tracking Research and Applications* (Stuttgart, Germany) *(ETRA '20 Short Papers)*. Association for Computing Machinery, New York, NY, USA, Article 25, 5 pages. https://doi.org/10.1145/3379156.3391336

[81] Asma Shakil, Christof Lutteroth, and Gerald Weber. 2019. CodeGazer: Making Code Navigation Easy and Natural With Gaze Input. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300306

[82] Ludwig Sidenmark, Christopher Clarke, Xuesong Zhang, Jenny Phu, and Hans Gellersen. 2020. Outline Pursuits: Gaze-Assisted Selection of Occluded Objects in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376438

[83] Ludwig Sidenmark and Hans Gellersen. 2019. Eye&Head: Synergetic Eye and Head Movement for Gaze Pointing and Selection. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 1161–1174. https://doi.org/10.1145/3332165.3347921

[84] Ludwig Sidenmark, Diako Mardanbegi, Argenis Ramirez Gomez, Christopher Clarke, and Hans Gellersen. 2020. BimodalGaze: Seamlessly Refined Pointing with Gaze and Filtered Gestural Head Movement. In *ACM Symposium on Eye Tracking Research and Applications* (Stuttgart, Germany) *(ETRA '20 Full Papers)*. Association for Computing Machinery, New York, NY, USA, Article 8, 9 pages. https://doi.org/10.1145/3379155.3391312

[85] Martin F. Stoelen and David L. Akin. 2010. Assessment of Fitts' law for quantifying combined rotational and translational movements. *Human factors* 52, 1 (2010), 63–77.

[86] Tableau. 2020. Tableaucolors. https://public.tableau.com/views/TableauColors/ColorPaletteswithRGBValues?%3Aembed=y&%3AshowVizHome=no&%3Adisplay_count=y&%3Adisplay_static_image=y

[87] Robert J. Teather and Wolfgang Stuerzlinger. 2014. Visual Aids in 3D Point Selection Experiments. In *Proceedings of the 2nd ACM Symposium on Spatial User Interaction* (Honolulu, Hawaii, USA) *(SUI '14)*. Association for Computing Machinery, New York, NY, USA, 127–136. https://doi.org/10.1145/2659766.2659770

[88] Anne M. Treisman and Garry Gelade. 1980. A feature-integration theory of attention. *Cognitive Psychology* 12, 1 (1980), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

[89] Mario H. Urbina and Anke Huckauf. 2007. Dwell time free eye typing approaches. In *Proceedings of the 3rd Conference on Communication by Gaze Interaction* (Leicester, UK) *(COGAIN '07)*. Association for Computing Machinery, New York, NY, USA.

[90] Boris Velichkovsky, Andreas Sprenger, and Pieter Unema. 1997. Towards gaze-mediated interaction: Collecting solutions of the "Midas touch problem". In *Human-Computer Interaction INTERACT'97*. Springer, 509–516.

[91] Sebastian Vetter, Jennifer Bützler, Nicole Jochems, and Christopher M Schlick. 2011. Fitts' law in bivariate pointing on large touch screens: Age-differentiated analysis of motion angle effects on movement times and error rates. In *International Conference on Universal Access in Human-Computer Interaction*. Springer, 620–628.

[92] Oleg Špakov and Darius Miniotas. 2005. Gaze-Based Selection of Standard-Size Menu Items. In *Proceedings of the 7th International Conference on Multimodal Interfaces* (Torento, Italy) *(ICMI '05)*. Association for Computing Machinery, New York, NY, USA, 124–128. https://doi.org/10.1145/1088463.1088486

[93] Colin Ware and Harutune H. Mikaelian. 1986. An Evaluation of an Eye Tracker as a Device for Computer Input2. *SIGCHI Bull.* 17, SI (May 1986), 183–188. https://doi.org/10.1145/30851.275627

[94] 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)*. ACM, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[95] Chia-Chien Wu, Oh-Sang Kwon, and Eileen Kowler. 2010. Fitts's Law and speed/accuracy trade-offs during sequences of saccades: Implications for strategies of saccadic planning. *Vision Research* 50, 21 (2010), 2142 – 2157. https://doi.org/10.1016/j.visres.2010.08.008

[96] Chun Yu, Yizheng Gu, Zhican Yang, Xin Yi, Hengliang Luo, and Yuanchun Shi. 2017. Tap, Dwell or Gesture? Exploring Head-Based Text Entry Techniques for HMDs. 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, 4479–4488. https://doi.org/10.1145/3025453.3025964

[97] Xuan Zhang and I. Scott MacKenzie. 2007. Evaluating eye tracking with ISO 9241-Part 9. In *International Conference on Human-Computer Interaction*. Springer, 779–788.