Direct vs. Score-based Selection: Understanding the Heisenberg Effect in Target Acquisition across Input Modalities in Virtual Reality

Category: Research

Figure 1: Our user study (N = 24) examines how direct and score-based selection methods with different input modalities (i.e., controller and hand inputs) are influenced by the Heisenberg Effect on task performance and user experience.

Abstract

Target selection is a fundamental interaction in virtual reality (VR). But the act of confirming a selection, such as a button press or pinch, can disturb the tracked pose and shift the intended target, which is referred to as the Heisenberg Effect. Prior research has mainly investigated controller input. However, it remains unclear how the effect manifests in the bare-hand input and how score-based techniques may mitigate the effect in different spatial variations. To fill the gap, we conduct a within-subject study to examine the Heisenberg Effect across two input modalities (i.e., controller and hand) and two selection mechanisms (i.e., direct and score-based). Our results show that hand input is more susceptible to the Heisenberg Effect, with direct selection more influenced by target width and score-based selection more sensitive to target density. Based on previous vote-oriented technique and our temporal analysis, we introduce weighted VOTE, a history-based intention accuracy model for target voting, that reweights recent interaction intent to counteract input disturbances. Our evaluation shows the method improves selection accuracy compared to baseline techniques. Finally, we discuss future directions for adaptive selection methods.

Index Terms: Virtual Reality, Input Modality, Heisenberg Effect, Score-based Selection, Interaction Techniques

1 Introduction

Target selection, as a fundamental interaction task in VR and augmented reality (AR), underpins a wide range of applications such as object manipulation and task execution. In current commercial VR systems, users perform selection tasks using various input modalities [24, 32], with handheld controllers and bare-hand tracking being the most prevalent due to their intuitiveness and versatility [17]. However, both input modalities involve the tracking of discrete actions like button presses or finger pinches, which can disturb the user's acquisition of the target and introduce unintended spatial offsets [4, 42]. This phenomenon, known as the **Heisenberg Effect** in spatial computing, describes how the discrete action of selection interferes with the target being observed or targeted, thereby affecting selection performance and user experience [42].

While the Heisenberg Effect has been systematically evaluated in direct controller tracking, several studies have informally noted that this effect may be more pronounced during hand-based interactions, likely due to greater movement variability and the lack of physical constraints [10, 19, 41]. However, the impact of this effect on bare-hand input has yet to be formally and quantitatively evaluated. Beyond input modalities, recent works have adopted score-based selection, such as snap-to raycasting [6, 12], focal selection [18, 41] and bubble-based mechanism [16, 46] as a way to mitigate the Heisenberg Effect. These score-based selection methods estimate user intent dynamically, reducing the need for precise alignment and avoiding the risk of selecting empty targets [6], reducing the extent to which target size affects performance as predicted by Fitts's law [9]. Unlike direct selections, score-based approaches greatly reduce the likelihood of void selections. As a result, when errors occur, they are more likely to cause irreversible or destructive actions immediately. Despite the widespread use of such methods, the influence of the Heisenberg Effect on the performance of score-based selections in VR remains under-examined, and there is a paucity of research focusing on the dense and occluded settings [20, 38, 45]. This gap motivates us to investigate how and to what extent the Heisenberg effect varies across different input modalities and mechanisms.

In this work, we first extend the analysis of the Heisenberg effect in target selection to both controller and hand tracking with both direct and score-based mechanisms. It is worth noting that the pinch, the most commonly used gesture to select distant objects [48, 27], is adopted as the confirmation action in our experiment. We found that during ballistic selections, the Heisenberg effect accounts for 80.57% of the selection errors for direct controller selection, 47.50% for score-based controller selection, 71.86% for direct hand selection, and 46.86% for score-based hand selection.

From the perspective of the Heisenberg Effect, our results provide a new explanation for why the selection accuracy of hand tracking is not as good as that of controller tracking. We also observe that direct selection is more sensitive to target size, while score-based mechanisms can partially compensate for the Heisenberg effect. However, their effectiveness varies significantly depending on target density (described as *target spacing* in our experiment). Furthermore, we find that as task difficulty increases, the Heisenberg Magnitude decreases, which may be attributed to users subconsciously stabilizing their hands in response to perceived difficulty, which reduces unintended movements.

To track intention shifts and accuracy trends across different selection techniques over time, we fit a third-degree polynomial to model the probability of correctly matching the intended object to the target (Section 5.1). Our evaluations indicate that although both direct and score-based hand selection exhibit lower final selection accuracy, the temporal records reveal intermittent moments of high precision. Building on these insights, and classic history-based selection techniques such as VOTE [25] and BackTracer [5], our quantitative analysis suggests that historical data should not contribute equally to final selection decisions. Accordingly, we refine the fitted temporal accuracy function into VOTE by incorporating a weighted function trained on our collected data. Our ablation study reveals that all input methods benefit from the weighted VOTE method, with the most pronounced improvement observed in score-based hand selection. The selection error rate is reduced substantially, from 21.96% to 7.54%, reaching a level comparable to controller input, where the score-based controller selections achieve an 8.18% error rate after improvement. We further analyzed individual intention accuracy histories to guide personalized adjustments. Although no significant improvement was observed, we outline implications for future adaptive optimization in VR. This study offers the following contributions:

- We extend the systematic analysis of the Heisenberg Effect in VR to different input modalities and selection mechanisms.
- We analyze the temporal distribution of users' intention shift and accuracy variation in selection events.
- We introduce and evaluate a weighted VOTE method to mitigate the Heisenberg effect by leveraging historical intention.

2 Related Work

To develop a general understanding of the Heisenberg Effect, we first review prior work that attempts to mitigate the effect and its implications across input modalities (Section 2.1). We then place particular emphasis on interaction techniques that leverage users' intention history, a line of research demonstrated to be especially efficient and directly informing our approach (Section 2.2).

2.1 Heisenberg Effect in Spatial Interaction

The Heisenberg effect, first noted by Bowman et al. [4] as a side effect of manipulating spatially tracked input devices (STIDs), was further analyzed in depth by Wolf et al. [42] for controller tracking in the target selection task. Through a systematic analysis of user behaviors, they identified key contributing factors that affect this effect (i.e., target size), and defined the Heisenberg Error as *"starting a selection inside the target but ending outside due to a disturbance of the input device."* Their investigations also outlined that discrete inputs, such as a button press, can physically disturb the STID, shifting the selection point during ray-cast interaction. While the Heisenberg effect is a well-explored concern for direct controller selection [5, 41, 43], its impact on bare-hand tracking remains largely unevaluated and bare-hand tracking is also prone to disturbances from discrete actions like pinch [10, 19].

Furthermore, from the perspective of input modality, controller and bare-hand are the two most widely used for object selection and manipulation in VR [11]. Controllers typically feature joysticks and buttons [23], offering benefits such as physical haptic feedback, ergonomics, and precise, consistent control [17]. Meanwhile, hand tracking often relies on vision-based algorithms [30] to detect gestures, enabling natural, device-free interaction that enhances users' sense of presence and body ownership [21, 15]. However, despite its intuitiveness, bare-hand input consistently underperforms in target selections compared to controllers in terms of selection efficiency and accuracy [21, 34, 41], especially for distant object interactions [33, 35]. Prior studies primarily attribute this inferiority to the limited reliability and precision of tracking algorithms [22, 2].

In this work, we aim to investigate the role of the Heisenberg effect as an overlooked factor resulting in performance degradation. We conduct a comparative quantitative analysis across VR input modalities to determine the extent to which it compromise selection performance and user experience.

2.2 Score-based Selections in VR

Besides techniques based on classic ray-casting, score-based selection techniques have been developed to improve target acquisitions in VR, particularly for interacting with distant or dynamic objects [12, 46, 29]. These methods assign an intention score for each potential target and snap to the user's desired one. Early approaches like IntenSelect [6, 12] and StickyRay [39] compute such a score based on the angle between the user's ray direction and targets. This principle has been incorporated into practical development workflows (e.g., Meta's *Distance Hand Grab Interactor* [27]). A key extension of this concept is focal-based selections [18, 41], which integrates spatial distance into the intent model and adapts it across various input modalities. However, a significant limitation of these score-based techniques in dense environments has been noted by FocalSelect [41] and BubbleRay [16]. Our experimental results corroborate this and the Heisenberg Effect may inadvertently destabilize selections, making it harder to acquire the intended target.

In addition to the snap-to heuristics, history-based scoring has emerged as another line of score-based techniques in VR to infer users' selection intent. For example, VOTE [25] selects the object with the highest occurrence within a fixed-time cache window. BackTracer [5] enhances this approach by dynamically adjusting the cache duration based on the perceived stability of targets. To mitigate the Heisenberg Effect, Wolf et al. [42] have proposed to leverage historical interaction traces to infer the originally intended target, compensating for spatial or temporal disruptions. However, this method still appears to increase many false positive selection errors, making it not robust in many cases.

Our research builds upon these efforts by incorporating weighted historical temporal inputs. Through quantitative analysis of user data in Section 3, we fitted functions to modulate the temporal contributions of historical inputs. These functions serve as a mitigation strategy against the Heisenberg Effect and are evaluated through an ablation study, demonstrating their effectiveness across various input modalities and selection mechanisms.

3 Experiment

In this section, we investigated how and to what extent various input modalities and selection mechanisms are influenced by the Heisenberg effect during target selection tasks. We further analyzed pre-selection user behavior and intention history.

3.1 Participants

We recruited 24 participants (11 females and 13 males), with an average age 25.0 (SD = 2.74). All individuals self-identified as right-handed and exhibited unimpaired vision and color discrimination

Figure 2: An example condition from our experimental design in Unity, which includes nine combinations of target size and spacing. In this example, the target width is 42 cm and the target spacing is 70 cm. The distance between the user and targets is fixed at 8 meters.

abilities. 15 of them had experience within VR as designers or developers, while others have less or no experience with VR devices. This study was approved by the Institutional Ethical Review Board (IRB), and informed consent was obtained from all participants.

3.2 Apparatus

The experiments were developed in Unity 2022 LTS, ran on Quest 3, and driven by a Windows 10 desktop (CPU: AMD Ryzen 7 5800H, GPU: Nvidia GeForce RTX 3060). Bare-hand and controller tracking were supported with the Meta Oculus SDK.

To ensure the fairness of our experiment, we implemented a unified white ray as the visual indicator for direct ray-casting tracking, which was supported for both hand and controller inputs by the Oculus SDK's ray interaction framework [29]. Following Wolf et al. [42], users performed target selection via the trigger button on the Quest 3 controller. The Oculus SDK registers the trigger's pressure on a continuous scale from 0 (not pressed) to 1 (fully pressed), with selection events triggered upon a full press. For bare-hand input, confirmation was triggered via thumb–index pinch gestures, detected by the built-in pose recognition system [28]. Details on detecting the start of a pinch action are provided in Section 3.3.2.

Based on the forward direction, we implemented Intenselect [6], a classical snap-to selection technique in VR, to support the score-based selection in two modalities. All score-based methods used a unified snap-to ray as the visual indicator and followed a spatiotemporal score to compute object scores, with the top-ranked object [6] selected. The scoring function is defined as follows:

$$s_t = s_{t-1}\gamma + \left(1 - \frac{\eta(t)}{\varepsilon}\right)(1-\gamma) \tag{1}$$

where s_t denotes the score at time t, $\eta(t)$ is the angle between the ray direction and the vector from the ray origin to the target's center at time t, ε is the angular threshold, and γ is the temporal decay factor, fixed at 0.5.

3.3 Testing Variables

3.3.1 Independent Variables

As shown in Figure 1, we evaluated four combinations of input modality and selection mechanism as **Input Techniques** to analyze the Heisenberg Effect: *(a) direct selection with controller tracking (DC)*, *(b) direct selection with hand tracking (DH)*, *(c) score-based selection with controller tracking (SC)*, and *(d) score-based selection with hand tracking (SH)*. Following the experimental setup in Wolf et al. [42], we introduced two environmental variables: **Target Width** *(14, 28, 42 cm)* and **Target Spacing** between the targets *(30, 50, 70 cm)*. We fixed the user's distance from the selection targets at 8 meters, which allowed us to measure the spatial shift of ray-casting using an offset angle relative to the participant. We also show the example experiment setup is shown in Figure 2.

3.3.2 Dependent Variables

To quantify the impact of the Heisenberg effect on user performance of target selection, we collected and evaluated the following dependent variables: **Selection Time**, **Overall Error**, **Heisenberg Error**, and **Heisenberg Magnitude**.

- The **Selection Time** was defined as the time taken to complete a single selection trial.
- The **Overall Error** was computed as the percentage of total selections that resulted in missed or incorrect targets.
- The **Heisenberg Error** captured the proportion of trials where the selection began inside the target area but concluded with a miss or incorrect selection.
- The **Heisenberg Magnitude** was measured as the offset in angular degrees between the start of the selection (start of button press or pinch gestures) and the end of the selection (end of button press or pinch gestures).

To accurately detect the start of a pinch gesture, we used two key signals: the frame-by-frame rate of changes in distance between the thumb and index finger (*velocityTI*) and the palm's rotational velocity (*velocityRot*) over a preceding time window (0.4s). The detection process involves two steps:

- **Identify the potential pinch onset:** We first find the initial moment when *velocityTI* exceeds a threshold of $0.05m/s$ and stays above it for at least 3 consecutive frames.
- **Determine the pinch start time:** From the identified onset, we search backward to locate the point where *velocityRot* has been continuously below $0.1deg/s$ for at least 5 frames. This time index is then identified as the exact start of a pinch action.

3.4 Procedure

We adopted a within-subjects design with a $4 \times 3 \times 3$ factorial structure: 4 **TECHNIQUE** conditions, 3 **TARGET WIDTH** levels, and 3 **TARGET SPACING** levels, following the grid layout from Lu et al. [16]. The order of TECHNIQUE conditions was counterbalanced with a Latin square to reduce order effects. The combinations of TARGET SIZE and TARGET SPACING were randomized across participants. Each condition was repeated three times to mitigate potential learning effects. In each scene associated with a specific TECHNIQUE condition, 49 (7×7 grid) spherical targets were displayed in a grid layout at a fixed distance of 8 meters from the participant. The spheres were colored either red or blue, with the intended selection target highlighted dynamically with a green border (Figure 2). Only the 25 inner targets (5×5 grid, excluding boundary targets) were used for selection tasks to ensure each target had four adjacent neighbors and a consistent selection difficulty.

For each TECHNIQUE condition in a single trial, participants were instructed to sequentially select 6 red-colored targets, which were randomly chosen from the 5×5 inner grid. Before the formal experiment, participants received instructions and were given sufficient time (approximately 2 minutes per technique) to practice. They could verbally indicate when they felt ready or ask questions if they encountered any confusion. During the formal sessions, rest periods were provided between trials, and participants were encouraged to pause or stop if they experienced fatigue. All participants completed the experiment while seated in a chair with armrests for

Figure 3: The performance summary of each technology under different target size conditions, including selection time, overall error, Heisenberg Error and Heisenberg Magnitude. The error bar represents the standard error.

comfort and arm support. The entire session lasted approximately 40 to 50 minutes per participant, including a semi-structured post-experiment interview where participants reflected on their experiences with the different selection techniques across conditions.

4 Result

In total, 2592 trials, including 21621 selection data points, were collected. We initially removed 32 trials (1.24% of 2592) where the completion time exceeded three standard deviations from the mean for each setup. These outliers were all due to some participants requesting short breaks and the oral reports accidentally triggering a selection event upon entering a new scene. The Shapiro-Wilk test indicated that the collected data regarding selection time were not normally distributed, leading us to conduct a repeated-measures ANOVA using Aligned Rank Transform [8]. Post-hoc pairwise comparisons were performed using paired t-tests with Bonferroni correction. The primary performance metric under different target widths is shown in Figure 3, while the comparison under different target spacing is shown in Figure 4. We denote statistical significance with * for $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$. The following analysis highlights significant results from the comparison under different spatial variations.

4.1 Quantitative Results

4.1.1 Selection Time

The effect of **TARGET WIDTH** on selection time for each input technique is shown in Figure 3. For DH, an RM-ANOVA revealed a significant effect of target size ($F_{2,609.06} = 87.856$, $p < 0.001$, $\eta^2 = 0.22$), with the small targets taking significantly longer than both medium ($t_{609} = 10.719$, $p < 0.001$) and large targets ($t_{609} = 12.108$, $p < 0.001$). For SH, target size also had a significant effect ($F_{2,610.8} = 5.366$, $p < 0.01$, $\eta^2 = 0.02$), where the large targets yielded shorter selection times than medium ($t_{610} = 2.892$, $p < 0.05$) and small targets ($t_{610} = 2.782$, $p < 0.05$). The difference between small and medium targets, as well as between medium and large targets, was not significant. For DC, a significant effect was observed ($F_{2,618.01} = 84.866$, $p < 0.001$, $\eta^2 = 0.22$), with small targets producing longer selection time than both medium ($t_{618} = 8.976$, $p < 0.001$) and large targets ($t_{618} = 12.632$, $p < 0.001$). Medium poses longer selection time than large targets ($t_{618} = 3.636$, $p < 0.001$). For SC, the effect of target size was not significant ($F_{2,615.02} = 2.2264$, $p = 0.109$), indicating stable selection performance across all target sizes.

The effect of **TARGET SPACING** on selection time for each input technique is shown in Figure 4. For DH, an RM-ANOVA revealed a significant effect of target spacing ($F_{2,609.08} = 3.371$, $p < 0.05$,, $\eta^2 = 0.01$). Only the spacing in near yielded longer selection time than the spacing in mid ($t_{611} = 2.596$, $p < 0.05$). For SH, target spacing had a significant effect ($F_{2,610.1} = 45.874$, $p < 0.001$, $\eta^2 = 0.13$), the spacing in near yielded longer selection time than the spacing in mid ($t_{610} = 7.134$, $p < 0.001$) and far ($t_{610} = 9.105$, $p < 0.001$). For DC, an RM-ANOVA revealed no significant effect of target spacing ($F_{2,618.02} = 5.021$, $p < 0.01$, $\eta^2 = 0.02$). Only spacing in mid yielded shorter selection time than the spacing in far ($t_{618} = -3.122$, $p < 0.01$). For SC, target spacing had a significant effect ($F_{2,615.02} = 10.063$, $p < 0.001$, $\eta^2 = 0.03$), where the spacing in near yielded longer selection time than the spacing in mid ($t_{609} = 4.356$, $p < 0.01$) and far ($t_{609} = 3.098$, $p < 0.01$).

4.1.2 Overall Error

The effect of **TARGET WIDTH** on Overall Error for each input technique is shown in Figure 3. For DH, an RM-ANOVA revealed a significant effect of target size ($F_{2,609.06} = 166.4$, $p < 0.001$, $\eta^2 = 0.35$), with the small targets posing significantly higher error rate than both medium ($t_{609} = 12.678$, $p < 0.001$) and large targets ($t_{609} = 17.695$, $p < 0.001$), also the medium targets posing higher error than large targets ($t_{609} = 4.998$, $p < 0.001$). For SH, target size reveals no significant effect ($F_{2,610.17} = 1.334$, $p = 0.2642$). For DC, a significant effect was observed ($F_{2,618.04} = 93.736$, $p < 0.001$, $\eta^2 = 0.23$), with small targets producing higher selection error than both medium ($t_{618} = 11.847$, $p < 0.001$) and large targets ($t_{618} = 11.863$, $p < 0.001$). For SC, the effect of target size was not significant ($F_{2,615.04} = 0.424$, $p = 0.655$), indicating stable performance across all target sizes.

Table 1: The percentage of the Heisenberg Error in the Overall Error in different spatial variations

	The proportion of the Heisenberg Error in Overall Error(%)						
	Target Radius			Target Spacing			
Input	Small (7cm)	Medium (14cm)	Large (21cm)	Near (30cm)	Mid (50cm)	Far (70cm)	Overall
DC	79.84	75.51	88.87	76.78	79.44	84.14	80.57
SC	43.48	53.51	44.89	54.82	46.18	28.31	47.50
DH	69.30	75.92	72.17	72.47	69.76	73.18	71.86
SH	48.10	44.43	47.94	43.14	51.58	46.35	46.86

Figure 4: The performance summary of each technology under different target spacing conditions, including selection time, overall error, Heisenberg Error and Heisenberg Magnitude. The error bar represents the standard error.

Table 2: Direction distribution of Heisenberg Magnitude

	Shift Direction (%)			
Input	TopRight	TopLeft	BottomRight	BottomLeft
DC	39.89	28.03	16.83	15.11
SC	47.41	29.89	10.75	11.31
DH	29.68	31.60	25.04	13.69
SH	34.90	33.31	21.64	10.14

The effect of **TARGET SPACING** on Overall Error for each input technique is shown in Figure 4. For DH, an RM-ANOVA revealed no significant effect of target spacing ($F_{2,609.09} = 1.201$, $p = 0.302$). For SH, target spacing had a significant effect ($F_{2,610.24} = 36.02$, $p < 0.001$, $\eta^2 = 0.11$), where the spacing in near posed higher selection error than spacing in mid ($t_{610} = 3.557$, $p < 0.01$) and far ($t_{610} = 4.901$, $p < 0.001$). The spacing in mid produced a higher selection error than spacing in far($t_{610} = 8.452$, $p < 0.001$). The spacing in mid posed a higher selection error than the spacing in far ($t_{610} = 4.901$, $p < 0.001$). For DC, an RM-ANOVA revealed a significant effect of target spacing ($F_{2,618.05} = 3.854$, $p < 0.05$, $\eta^2 = 0.01$). The spacing in near posed a lower error than the spacing in far ($t_{618} = -2.578$, $p < 0.05$). For SC, target spacing had a significant effect ($F_{2,615.04} = 30.389$, $p < 0.001$, $\eta^2 = 0.09$), where the spacing in near posed a higher selection error than spacing in mid ($t_{615} = 5.407$, $p < 0.001$) and far ($t_{615} = 7.562$, $p < 0.001$).

4.1.3 Heisenberg Error

The effect of **TARGET WIDTH** on Overall Error for each input technique is shown in Figure 3. For DH, an RM-ANOVA revealed a significant effect of target size ($F_{2,609.12} = 92.737$, $p < 0.001$, $\eta^2 = 0.23$), with the small targets posing significantly higher Heisenberg Error than both medium ($t_{609} = 8.428$, $p < 0.001$) and large targets ($t_{609} = 13.477$, $p < 0.001$), also the medium targets posing higher Heisenberg Error than large targets ($t_{609} = 5.033$, $p < 0.001$). For SH, target size reveals no significant effect ($F_{2,610.33} = 1.477$, $p = 0.229$). For DC, a significant target size effect was observed ($F_{2,618.05} = 71.457$, $p < 0.001$, $\eta^2 = 0.19$), with small targets producing higher Heisenberg Error than both medium ($t_{618} = 10.744$, $p < 0.001$) and large targets ($t_{618} = 9.907$, $p < 0.001$). For SC, the effect of target size was not significant ($F_{2,615.09} = 1.885$, $p = 0.15$).

The effect of **TARGET SPACING** on Overall Error for each input technique is shown in Figure 4. For DH, an RM-ANOVA revealed no significant effect of target spacing ($F_{2,609.16} = 2.594$, $p = 0.0756$). For SH, target spacing had a significant effect ($F_{2,610.47} = 14.362$, $p < 0.001$, $\eta^2 = 0.04$), where the spacing in near yielded a higher Heisenberg Error than spacing in far ($t_{610} = 4.520$, $p < 0.001$). The spacing in mid also posed a higher Heisenberg Error than the spacing in far ($t_{610} = 4.754$, $p < 0.001$). For DC, an RM-ANOVA revealed a significant effect of target spacing with a small effect size ($F_{2,618.06} = 5.170$, $p < 0.01$, $\eta^2 = 0.02$). The

spacing in near produced a lower Heisenberg Error than the spacing in far ($t_{618} = -3.148$, $p < 0.01$). For SC, target spacing had a significant effect ($F_{2,615.08} = 30.534$, $p < 0.001$, $\eta^2 = 0.09$), where the spacing in near yielded a higher Heisenberg Error than spacing in mid ($t_{615} = 5.013$, $p < 0.001$) and spacing in far ($t_{615} = 7.672$, $p < 0.001$). The spacing in mid also posed a higher Heisenberg Error than spacing in far ($t_{615} = 2.660$, $p < 0.05$).

4.1.4 Heisenberg Magnitude

The effect of **TARGET WIDTH** on Heisenberg Magnitude for all techniques is shown in Figure 3. For DH, an RM-ANOVA revealed a significant effect of target size ($F_{2,609.07} = 24.241$, $p < 0.001$, $\eta^2 = 0.07$), with the small targets produced lower Heisenberg Magnitude than both medium ($t_{609} = -4.053$, $p < 0.001$) and large targets ($t_{609} = -6.929$, $p < 0.001$). The medium objects produced lower Heisenberg Magnitude than the large objects ($t_{609} = -2.868$, $p < 0.05$). For SH, target size had no significant effect ($F_{2,610.09} = 0.184$, $p = 0.832$). For DC, a significant effect was observed ($F_{2,618.03} = 43.178$, $p < 0.001$, $\eta^2 = 0.12$), with small targets posing higher Heisenberg Error than both medium ($t_{618} = -6.830$, $p < 0.001$) and large targets ($t_{618} = -8.870$, $p < 0.001$). For SC, target size had no significant effect ($F_{2,615.01} = 1.686$, $p = 0.186$).

The effect of **TARGET SPACING** on Heisenberg Magnitude for each input technique is shown in Figure 4. For DH, an RM-ANOVA revealed a significant effect of target spacing ($F_{2,609.07} = 26.328$, $p < 0.001$, $\eta^2 = 0.08$), with the spacing in near produced lower Heisenberg Magnitude than both spacing in mid ($t_{609} = -5.330$, $p < 0.001$) and far ($t_{609} = -6.928$, $p < 0.001$). For SH, target spacing had a significant effect ($F_{2,610.1} = 82.482$, $p < 0.001$, $\eta^2 = 0.21$) with a large effect size, with the spacing in near produced lower Heisenberg Magnitude than both spacing in mid ($t_{610} = -9.251$, $p < 0.001$) and far ($t_{610} = -12.344$, $p < 0.001$). The spacing in mid also produced a lower Heisenberg Magnitude than the spacing in far ($t_{610} = -3.102$, $p < 0.01$). For DC, target spacing had a significant effect ($F_{2,618.03} = 34.877$, $p < 0.001$, $\eta^2 = 0.10$), with the spacing in near produced lower Heisenberg Magnitude than both spacing in mid ($t_{618} = -6.517$, $p < 0.001$) and far ($t_{618} = -7.780$, $p < 0.001$). For SC, target spacing had a significant effect ($F_{2,615.01} = 38.674$, $p < 0.001$, $\eta^2 = 0.11$), with the spacing in near produced lower Heisenberg Magnitude than both spacing in mid ($t_{615} = -4.982$, $p < 0.001$) and far ($t_{615} = -8.766$, $p < 0.001$). The spacing in mid also produced a lower Heisenberg Magnitude than the spacing in far ($t_{615} = -3.758$, $p < 0.01$).

4.1.5 Directional Bias with Heisenberg Effect

To further quantify the systematic directional offset introduced by the Heisenberg Effect, we analyzed the spatial displacement for each selection. This offset was calculated as the angular distance between the aiming vector at the initiation of a discrete action (i.e., press or pinch) and the vector at its completion. Each selection's final position was categorized relative to the initial aiming vector into one of four quadrants: Top, Bottom, Left, or Right. The aggregated distribution of these directional offsets is presented in Tab. 2.

Consistent with the findings of Wolf et al [42], our results reveal a distinct directional bias for controller-based interactions. We observed a consistent upward shift in the final selection point for 72.6% of trials using the Quest 3 controller. We attribute this systematic error to the physical location and actuation mechanics of the trigger button, which tends to cause a slight upward rotation of the controller during a press. In contrast, the spatial offset for bare-hand input showed no dominant directional bias and was significantly more distributed. We hypothesize that this variance stems from the inherent diversity in individual hand physiology and the unique pinch gestures adopted by each participant, leading to less predictable displacement vectors upon selection.

4.2 Evaluation and Discussion

4.2.1 General Characteristics of the Heisenberg Effect

The quantitative analysis of Heisenberg Magnitude shows that for direct selection, the Heisenberg Magnitude decreases as the target width decreases. Prior work attributed this phenomenon to anticipatory postural adjustments that modulate muscle tension and arm posture depending on the perceived target size [26, 42]. Our study extends this understanding by demonstrating that, regardless of whether participants performed direct or score-based selection, the Heisenberg Magnitude also decreases as target spacing becomes smaller. This suggests that when task difficulty increases, participants reinforce motor control strategies to deal with higher selection error rates, thereby reducing the impact of the Heisenberg Effect. This interpretation is further supported by post-hoc interviews, in which the majority of participants (22/24) reported elevated nervousness when performing the most challenging tasks.

4.2.2 Bare-Hand vs. Controller Input

Consistent with prior research [42], direct controller input is subject to the Heisenberg Effect. Our findings indicate that the performance limitations of this effect are most pronounced only for very small targets, as we found no significant performance difference when selecting medium versus large objects.

Our analysis reveals that bare-hand input is more susceptible to the Heisenberg Effect than controller input. This is proven by a significantly larger Heisenberg Magnitude and a substantially higher Heisenberg Error for the hand condition. While the absolute number of Heisenberg Errors was greater for bare-hand input, they constituted a smaller proportion of total errors compared to the controller condition. This suggests that, in addition to the Heisenberg Effect, the initial pointing accuracy of bare-hand tracking before the selection gesture is lower than that of controller inputs.

We found that in the period before pinch action, slight head shaking or camera movement may also disturb hand tracking. Unlike controller inputs, which leverage IMU-based tracking for relatively more stable spatial localization, hand tracking relies solely on vision-based approaches and is more sensitive to camera motion. Consequently, even minor head movements can result in unexpected spatial shifts of hand input. Nevertheless, following the definition of the Heisenberg Effect established in prior work [42], we did not categorize this case as part of the Heisenberg Error.

4.2.3 Direct vs. Score-based Selection

From the quantitative analysis of the above performance metrics, it can be concluded that direct selection is more sensitive to changes in target width, but its performance only drops drastically when the targets are small. In contrast, score-based selection avoids null selection and thus is not significantly affected by changes in target width. However, when other objects around the target are more densely packed, the performance of score-based selection is significantly impacted, showing a clear negative correlation. We also found a significant difference in Heisenberg Magnitude between DH and SH ($t_{1246} = -3.056$, $p < 0.01$) as well as DC and SC ($t_{1260} = -14.470$, $p < 0.001$). This means the score-based selection produces more Heisenberg Magnitude than the direct selection. Overall, score-based selection offers a noticeable performance improvement over direct selection, especially in the case of hand selection to reduce the selection error rate.

4.2.4 Visual Indicator Matters in the Heisenberg Effect

When some participants (7/24) felt that score-based selection was inferior to direct selection, they noted it to "*the sticky ray cannot determine the endpoint exactly*." Our observations support the counterintuitive viewpoint: during score-based selection, the scores of nearby objects often approach that of the intended target, making it difficult to distinguish. Although the curvature of the Bezier curve

Figure 5: Distribution of selection endpoints across input techniques with a target width of 28 cm and spacing of 70 cm. Endpoints under direct selection (a, c) are more concentrated around the targets, whereas score-based selection (b, d) shows greater dispersion.

can partly indicate the original direction of ray and its alignment with the target, participants still struggled to reliably control the ray to prevent unintended shifts.

Further evidence of the role of visual indicators in target selection is shown in Figure 5, showing the distribution of selection endpoints for different input techniques. Endpoints of direct selection cluster tightly around targets, whereas score-based selection produces a more scattered endpoint distribution. This contrast suggests that visual indicators are critical for stabilizing endpoint control. In direct selection, the clear spatial alignment between the ray and the target provides strong visual cues, enabling participants to converge on their intended objects. In score-based selection, the absence of such explicit spatial anchoring leads to higher endpoint variability, especially in high-density environments. These findings underscore that the presence and clarity of visual indicators influence both user performance and confidence in target selection.

4.2.5 Bias Caused by Different Devices and Setups

Compared to the quantitative results reported by Wolf et al. [42], our measured Heisenberg Magnitude for direct controller tracking (avg = 0.358°) is slightly lower than theirs (avg > 0.5°). We hypothesize that this discrepancy stems from hardware differences: our study employed the Meta Quest 3, a newer VR device with different button placements and pressure thresholds, whereas Wolf et al. used the HTC Vive. Additional testing is needed to evaluate how button position and device design influence directional shifts across different VR devices with controllers.

Our results indicate that the Heisenberg Effect accounts for 80.57% of overall errors in practical target selection tasks, markedly higher than the 30.49% reported in ballistic tasks [42], and closely aligned with errors observed in stationary selection tasks (81.98%). This discrepancy is likely due to our use of visual highlighting around the intended object, which reduced participants' initial positioning errors [3]. This suggests that the Heisenberg Effect is not only device-dependent but also shaped by the visual highlighting of intended targets and the feedback provided during selection.

5 Proposed Mitigation Method

5.1 Observation and Motivation on Intention History

We first rescale the time frames in the history cache for each selection record to a 0–1 scale based on their relative positions within the full event timeline. This normalization accounts for instances where the duration of a selection event lasted shorter than the 0.4-second threshold. We then split the collected data into a training set (19 participants, 17187 selection records) and a test set (5 participants, 4157 selection records). Observing that intention accuracy in the preceding time window follows an S-shaped trend, we model the likelihood of correctly matching the intended object to the target using a third-degree polynomial for all input methods. The formulas for normalization of time length and fitting function for intention accuracy are as follows:

$$\tilde{T} = \frac{t - T_c}{T_e} \tag{2}$$

$$f(\tilde{T}) = a\tilde{T}^3 + b\tilde{T}^2 + c\tilde{T} + d \tag{3}$$

where $\tilde{T}$ represents the relative time of a candidate item within the intention history for each selection. T_c denotes the timestamp at which the user performs the confirmation action, and t is the timestamp of the candidate item being evaluated. $T_e = 0.4s$ represents the expiration window that defines the valid temporal range for recording past interaction frames. The coefficients a, b, c, and d are empirically fitted based on the training data, and the overall fitting function is shown in the first row of Figure 6, demonstrating the varied selection accuracy in the interaction process.

Consistent with the multiple process model that includes target searching, aiming, and confirmation [44], intention accuracy for all methods first increases, then drops sharply during the final confirmation stage. As shown in Figure 6 b) and d), this indicates a strong influence of the Heisenberg Effect on selection performance, which was particularly pronounced in the DH and SH approaches (nearly **20% accuracy loss**). Notably, the score-based hand tracking achieved an accuracy of nearly 98%, which is comparable to the two controller conditions, indicating that the incorporation of historical information can significantly improve the selection accuracy by alleviating the Heisenberg effect.

However, a unified weighting function is insufficient. As shown in the second row of Figure 6, individual intention accuracy trends can deviate significantly from the overall model, even though most participants conform to the general pattern. This variability stems from both fine-grained factors (e.g., hand tremor, pinch speed) and coarse-grained factors (e.g., hand posture). Our findings suggest that common compensation mechanisms, which backtrack a selection to a fixed time frame as proposed in prior work [42], can be suboptimal. Such a direct backtracking method often increases false positive errors in direct selections (i.e., DC and DH). Our temporal analysis of intention accuracy indicates that historical data should be weighted differently to mitigate the Heisenberg Effect. Consequently, this motivates us to explore how to adjust the fitted function to accommodate individual behavioral variations.

Figure 6: Trends of user intention accuracy over time across four selection modalities and mechanisms. Each curve is fitted with a third-degree polynomial within 0.4s before selection. Relative time (x axis) is normalized across selection events, where **0.0 denotes selection completion**. Accuracy of Intention (y axis) is defined as the probability that the user's selection ray is pointing at or snapping to the correct target object, where **1.0 denotes that the intended object must be the target one.** The first row shows overall accuracy curves trained on data from P1–P19, illustrating general trends. The bottom row highlights individual participants whose accuracy patterns deviate from the overall trend.

5.2 Weighted VOTE Based on Empirical Data

To integrate our findings to improve the performance in target selection tasks, we referred to previous history-based techniques [25, 5] and implemented the classic VOTE methods as a baseline. Specifically, VOTE is a method to assign the object with the highest frequency in a specific time slot (0.4s in our implementation) to be the final selection and is formulated as follows:

$$O^* = \arg\max_{o \in \tilde{O}} \sum_{\tilde{T} \leq 1} \mathbf{1}[V(\tilde{T}) = o] \tag{4}$$

where O^* denotes the final object returned by Weighted VOTE. $\tilde{O}$ represents the finite set of all candidate (selectable) objects in the intention history. $V(\tilde{T})$ indicates the object indicated by the underlying interaction technique at time $\tilde{T}$, equals the special symbol "null" when no object is currently indicated. $1[\cdot]$ indicates a function returns 1 if the condition is true, 0 otherwise.

Considering our findings, **each time frame requires different voting weights**, and the higher the historical accuracy in a specific time frame, the higher the voting weights should be. In order to enhance the precision of moments and give them higher voting weights, we applied a sigmoid function to enhance the contrast of voting weights. The weighting function can be formulated as:

$$W(\tilde{T}) = \frac{1}{1 + e^{-k(f(\tilde{T}) - A)}} \tag{5}$$

$$O^* = \arg\max_{o \in \tilde{O}} \sum_{\tilde{T} \leq 1} W\left(\tilde{T}\right) \cdot \mathbf{1}[V(\tilde{T}) = o] \tag{6}$$

where the $W(\tilde{T})$ indicates a temporal weighting function that assigns a non-negative weight to each vote based on its relative time. k indicates an adjustable parameter controlling the steepness of the sigmoid weighting curve, assigned with 10 in our implementation. A denotes the shift parameter of the sigmoid. For each fitted weight function, it is adaptively defined as the empirical two-thirds quantile of the function values within the $[0, 1]$ interval.

To address individual differences, we developed the Adaptive Weighted VOTE for a more personalized backtrack strategy. Specifically, we fitted a cubic function using the first 20% of the test user's selection history and interpolated it with the overall fitted function (P1–P19), assigning a higher interpolation weight (i.e., 0.4 in our implementation) to the individual model. This approach allows the system to adapt more rapidly to personal behavior patterns while retaining the stability of the general model.

5.3 Evaluation on Mitigation Strategies

In addition to the original method, VOTE, and our proposed Weighted VOTE, we selected the temporal point corresponding to the peak of the third-degree polynomial fitted on the training data as the backtracking reference time frame. We refer to this approach as "Shift to Before" in our evaluations. After fitting the voting weights for each input method, we performed a comparative evaluation on both the original selection and the improved techniques. To ensure fair evaluation across all mitigation strategies, we use only the last 80% of the data in the test set for evaluation.

As shown in Tab. 3, Weighted VOTE significantly reduces the error rate across all input techniques, though its impact varies. The most substantial improvement was observed in SH, where applying Weighted VOTE reduced its error rate to a level comparable with controller tracking. This success stems from the longer aiming times typical of bare-hand interaction, which provides a richer history of user intention that VOTE effectively leverages. Conversely, Weighted VOTE provided minimal benefits for direct selection techniques (DH and DC). For these methods, historical data was already dominated by the correct target and numerous empty objects. Since VOTE is designed to ignore empty objects by default, its application offered no significant improvements. Such features of voting-based methods also explain why Adaptive Weighted VOTE did not demonstrate outstanding performance in direct selections. Only a minor performance improvement was found in the SH technique. We discuss adaptive techniques under VR in Section 6.2.

6 Discussion and Future Work

In this study, we conduct a systematic analysis of how the Heisenberg Effect matters with various input techniques and introduce a

Table 3: Evaluation of selection error with mitigation strategies in the test set. The bold indicates the technique with the lowest error rate.

	Error Rate in the Test Set (%)				
Input	Origin	Shift to Before	VOTE	Weighted VOTE	Adaptive Weighted VOTE
DC	15.20	12.06	3.49	2.96	**2.94**
SC	13.35	14.72	11.56	**8.18**	8.56
DH	32.92	25.00	10.42	**10.42**	10.65
SH	21.96	8.22	11.90	7.93	**7.54**

potential mitigation strategy based on intention data. This section demonstrates findings and implications for future work.

6.1 The Heisenberg Effect with Other Input Methods, Devices, and Visual Indicators

Our work fills the gap in the comparative analysis of bare-hand input and score-based selection techniques, which are widely used in VR development. While different input modalities and mechanisms are all subject to the Heisenberg Effect, the variables that drive its impact and the degree of influence vary across techniques. Through an empirical study, we disentangle the characteristics of the Heisenberg Effect under different inputs and propose a unified approach that can be applied to diverse interaction techniques.

Based on the distinct characteristics we observed across different input techniques, future work is encouraged to **supplement the analysis of the Heisenberg Effect on more input modalities**. For example, compound input modalities, such as Hand + Eye input, have become increasingly prevalent in the VR community [31, 37] and are now integrated into commercial systems such as the Apple Vision Pro [1]. The community has been actively exploring alternative or complementary modalities, including head-ray pointing and foot-based selection [40], to broaden interaction possibilities. Future work may follow our experimental framework and adopt both quantitative measures (i.e., Heisenberg Magnitude, selection time, error rates) and qualitative assessments (i.e., ergonomics and comfort) to capture their performance and usability characteristics.

Beyond algorithms and mechanisms for target selection, the difference in Heisenberg Magnitude between our results and those reported in prior work [4] provides evidence that the Heisenberg Effect may vary considerably across different hardware platforms, even when using similar controller inputs. This **highlights the importance of device-specific evaluation**. As new devices continue to emerge [7, 14], we encourage researchers and developers to measure the Heisenberg Effect as part of performance benchmarks to ensure consistent interaction quality across hardware ecosystems.

Our observation also reveals that the snap-to mechanism hindered precise ray-casting direction adjustments due to uncertainty about the endpoint's direction. This feedback indicates that the elevated error rate in dense environments is influenced by the visual indicator and feedback design. Our study exclusively employed a unified visual feedback design that combined snap-to raycasting with target highlighting. Whether alternative feedback mechanisms might affect selection accuracy, particularly in the presence of the Heisenberg Effect, requires further investigation. Previous studies have investigated how visual feedback influences selection performance in sparse or randomized settings [3], and have also explored approaches that shift the snap-to location from the object center to arbitrary points within the object [12]. Instead of the visualization of ray-casting, FocalSelect enabled a floating sphere in the occlusion plane to indicate the endpoint in score-based selection [41]. Future work should **explore a broader design space of visual indicators and feedback** and decouple the variables of visual indicators from score-based selection. For example, score-based selection paired with direct ray-casting (i.e., straight line indicator) could be effective in densely populated environments.

6.2 Implication for History-based Adaptive Techniques

Drawing on empirical observations and quantitative analysis, we propose a guideline that **leverages historical intention and introduces a data-driven mitigation strategy** to reduce the Heisenberg Effect. While our offline analysis confirmed the effectiveness of the weighted VOTE method in the test set, its impact on user experience in real-time interaction remains unclear. Future work may conduct qualitative user studies to assess its practical usability and implications [25]. Furthermore, our implementation could be enhanced by replacing the fixed time window with an adaptive cache duration responsive to environmental factors [5]. To leverage the contextual strengths of different selection techniques, Lai et al. [13] have proposed intelligent switching strategies that dynamically choose the optimal technique. Constrained by temporal voting and shifting, our compensation approach becomes less effective when the correct target is absent from the intention history. Future research could explore spatial shifting or spatiotemporal backtracing methods to further mitigate the Heisenberg Effect.

Furthermore, we observed significant variance in user selection behaviors, evidenced by the different intention correction probabilities between individually-fitted and generally-fitted models. The lack of significant improvement in our adaptive method likely stems from its exclusive reliance on personal temporal features. The fitting model neglected other crucial behavioral inputs, such as kinematic data from hand and eye tracking [47], which may have better accounted for these individual differences. In addition, VR and AR provide a comprehensive platform for capturing user behaviors. To enable personalized user experiences, such as the target selection task in our work, it is essential to design appropriate sampling and analysis methods while also considering the real costs in practical scenarios. In a single session, users may frequently switch between multiple applications, change between sitting and standing, and alter their hand postures. Each change in context or posture can make previously learned weights invalid, requiring recalibration. Future work could incorporate real-time error detection [36] to label the correctness of the intention history and enable dynamic, human-in-the-loop adaptation of the weighted VOTE technique.

7 Conclusion

In this work, we evaluate the Heisenberg Effect in target selection tasks across four widely used VR interaction techniques, focusing on their impact on selection performance. To assess the influence of environmental factors, we conducted a user study ($N = 24$) with nine combinations of target size $\times$ target spacing. Results indicate that direct ray-casting performs poorly with small targets, while the score-based technique is less effective in densely spaced environments. The Heisenberg Effect appears to be a primary factor underlying the inferior performance of bare-hand input compared to controllers. Building on insights from a temporal analysis of accuracy in historical intentions, we introduce a weighted VOTE method to enhance selection performance. Our evaluation demonstrates that it significantly improves the accuracy of score-based selections with bare-hand input, which is comparable to that of controller input.

References

[1] Apple Inc. Apple Vision Pro. https://www.apple.com/sg/apple-vision-pro/, 2023. Retrieved August 10, 2025. 9

[2] F. Argelaguet and C. Andujar. Special Section on Touching the 3rd Dimension: A survey of 3D object selection techniques for virtual environments. *Comput. Graph.*, 37(3):121–136, May 2013. doi: 10.1016/j.cag.2012.12.003 2

[3] M. Baloup, T. Pietrzak, and G. Casiez. RayCursor: A 3D Pointing Facilitation Technique based on Raycasting. In *Proceedings of the*

2019 CHI Conference on Human Factors in Computing Systems, CHI '19, p. 1–12. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290605.3300331 7, 9

[4] D. Bowman, C. Wingrave, J. Campbell, and V. Ly. Using Pinch Gloves for Both Natural and Abstract Interaction Techniques in Virtual Environments. *Proceedings of HCI International*, 01 2001. 1, 2, 9

[5] K. Chen, H. Wan, S. Zhao, and X. Liu. BackTracer: Improving Raycasting 3D Target Acquisition by Backtracking the Interaction History. *Int. J. Hum.-Comput. Stud.*, 176(C), Aug. 2023. doi: 10.1016/j.ijhcs.2023.103045 2, 8, 9

[6] G. de Haan, M. Koutek, and F. H. Post. IntenSelect: Using Dynamic Object Rating for Assisting 3D Object Selection. In *Proceedings of the 11th Eurographics Conference on Virtual Environments*, EGVE'05, p. 201–209. Eurographics Association, Goslar, DEU, 2005. 1, 2, 3

[7] T. Drey, J. Gugenheimer, J. Karlbauer, M. Milo, and E. Rukzio. VRSketchIn: Exploring the Design Space of Pen and Tablet Interaction for 3D Sketching in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, p. 1–14. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3313831.3376628 9

[8] L. A. Elkin, M. Kay, J. J. Higgins, and J. O. Wobbrock. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. In *The 34th Annual ACM Symposium on User Interface Software and Technology*, UIST '21, p. 754–768. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3472749.3474784 4

[9] P. M. Fitts. The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. *Journal of Experimental Psychology*, 47 6:381–91, 1954. 1

[10] J. Gabel, S. Schmidt, K. Pfeuffer, and F. Steinicke. Guiding Handrays in Virtual Reality: Comparison of Gaze- and Object-Based Assistive Raycast Redirection. In *Proceedings of the 2024 ACM Symposium on Spatial User Interaction*, SUI '24. Association for Computing Machinery, New York, NY, USA, 2024. doi: 10.1145/3677386.3682080 1, 2

[11] J. Kangas, S. K. Kumar, H. Mehtonen, J. Järnstedt, and R. Raisamo. Trade-Off between Task Accuracy, Task Completion Time and Naturalness for Direct Object Manipulation in Virtual Reality. *Multimodal Technologies and Interaction*, 6(1), 2022. doi: 10.3390/mti6010006 2

[12] M. Krüger, T. Gerrits, T. Römer, T. Kuhlen, and T. Weissker. IntenSelect+: Enhancing Score-Based Selection in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 30(5):2829–2838, 2024. doi: 10.1109/TVCG.2024.3372077 1, 2, 9

[13] C.-J. Lai, M. Sousa, T. Zhang, L. Sidenmark, and T. Grossman. Adaptique: Multi-objective and Context-Aware Online Adaptation of Selection Techniques in Virtual Reality. *arXiv preprint arXiv:2508.08505*, 2025. 9

[14] N. Li, T. Wu, Z. He, L. Shen, T. Luo, T. Han, B. Gao, Y. Zhang, L. Zhang, F. Tian, and Q. Wang. A Dual-Stick Controller for Enhancing Raycasting Interactions with Virtual Objects. In *2025 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 381–391, 2025. doi: 10.1109/VR59515.2025.00062 9

[15] L. Lin, A. Normoyle, A. Adkins, Y. Sun, A. Robb, Y. Ye, M. Di Luca, and S. Jörg. The Effect of Hand Size and Interaction Modality on the Virtual Hand Illusion. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 510–518, 2019. doi: 10.1109/VR.2019.8797787 2

[16] Y. Lu, C. Yu, and Y. Shi. Investigating Bubble Mechanism for Ray-Casting to Improve 3D Target Acquisition in Virtual Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 35–43, 2020. doi: 10.1109/VR46266.2020.00021 1, 2, 3

[17] T. Luong, Y. F. Cheng, M. Möbus, A. Fender, and C. Holz. Controllers or Bare Hands? A Controlled Evaluation of Input Techniques on Interaction Performance and Exertion in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 29(11):4633–4643, 2023. doi: 10.1109/TVCG.2023.3320211 1, 2

[18] J. Ma, J. Qian, T. Zhou, and J. Huang. FocalPoint: Adaptive Direct Manipulation for Selecting Small 3D Virtual Objects. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 7(1), Mar. 2023. doi: 10.1145/3580856 1, 2

[19] H. Mao, M. Gonzalez-Franco, V. Phadnis, E. J. Gonzalez, and I. Chatterjee. RestfulRaycast: Exploring Ergonomic Rigging and Joint Amplification for Precise Hand Ray Selection in XR. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference*, DIS '25, p. 28–39. Association for Computing Machinery, New York, NY, USA, 2025. doi: 10.1145/3715336.3735677 1, 2

[20] M. Maslych, Y. Hmaiti, R. Ghamandi, P. Leber, R. K. Kattoju, J. Belga, and J. J. LaViola. Toward Intuitive Acquisition of Occluded VR Objects Through an Interactive Disocclusion Mini-map. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 460–470, 2023. doi: 10.1109/VR55154.2023.00061 1

[21] A. Masurovsky, P. Chojecki, D. Runde, M. Lafci, D. Przewozny, and M. Gaebler. Controller-Free Hand Tracking for Grab-and-Place Tasks in Immersive Virtual Reality: Design Elements and Their Empirical Study. *Multimodal Technologies and Interaction*, 4(4), 2020. doi: 10.3390/mti4040091 2

[22] S. Mayer, V. Schwind, R. Schweigert, and N. Henze. The Effect of Offset Correction and Cursor on Mid-Air Pointing in Real and Virtual Environments. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, p. 1–13. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3174227 2

[23] D. Mendes, F. M. Caputo, A. Giachetti, A. Ferreira, and J. Jorge. A Survey on 3D Virtual Object Manipulation: From the Desktop to Immersive Virtual Environments. *Computer Graphics Forum*, 2019. doi: 10.1111/cgf.13390 2

[24] Meta Platforms, Inc. Quest 3. https://www.meta.com/quest/quest-3/, 2023. Retrieved August 10, 2025. 1

[25] A. G. Moore, J. G. Hatch, S. Kuehl, and R. P. McMahan. VOTE: A Ray-casting Study of Vote-oriented Technique Enhancements. *International Journal of Human-Computer Studies*, 120:36–48, 2018. doi: 10.1016/j.ijhcs.2018.07.003 2, 8, 9

[26] S. Nana-Ibrahim, S. Vieilledent, P. Leroyer, Y. Kerlirzin, and Y. Brenière. Target Size Modifies Anticipatory Postural Adjustments and Subsequent Elementary Arm Pointing. *Experimental Brain Research*, 184(2):255–260, 2008. doi: 10.1007/s00221-007-1178-7 6

[27] Oculus Developer Documentation. Distance Hand Grab Interactors. https://developer.oculus.com/documentation/unity/unity-isdk-distance-hand-grab-interaction/, 2023. Retrieved August 10, 2025. 1, 2

[28] Oculus Developer Documentation. Oculus Pose Detection, 2025. Retrieved August 10, 2025. 3

[29] Oculus Developer Documentation. Ray Interactions. https://developers.meta.com/horizon/documentation/unity/unity-isdk-ray-interaction, 2025. Retrieved August 10, 2025. 2, 3

[30] M. Oudah, A. Al-Naji, and J. Chahl. Hand Gesture Recognition Based on Computer Vision: A Review of Techniques. *Journal of Imaging*, 6(8), 2020. doi: 10.3390/jimaging6080073 2

[31] K. Pfeuffer, B. Mayer, D. Mardanbegi, and H. Gellersen. Gaze + pinch interaction in virtual reality. In *Proceedings of the 5th Symposium on Spatial User Interaction*, SUI '17, p. 99–108. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3131277.3132180 9

[32] PICO Immersive Pte.ltd. Pico 4. https://www.picoxr.com/global/products/pico4, 2023. Retrieved August 10, 2025. 1

[33] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa. The Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR. In *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology*, UIST '96, p. 79–80. Association for Computing Machinery, New York, NY, USA, 1996. doi: 10.1145/237091.237102 2

[34] H.-R. Rantamaa, J. Kangas, S. K. Kumar, H. Mehtonen, J. Järnstedt, and R. Raisamo. Comparison of a VR Stylus with a Controller, Hand Tracking, and a Mouse for Object Manipulation and Medical Marking Tasks in Virtual Reality. *Applied Sciences*, 13(4), 2023. doi: 10.3390/app13042251 2

[35] J. Schjerlund, K. Hornbæk, and J. Bergström. Ninja Hands: Using Many Hands to Improve Target Selection in VR. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21. Association for Computing Machinery, New York, NY, USA,

2021. doi: 10.1145/3411764.3445759 2

[36] N. Sendhilnathan, T. Zhang, D. Bethge, M. Nebeling, T. Grossman, and T. R. Jonker. A Multimodal Approach for Targeting Error Detection in Virtual Reality Using Implicit User Behavior. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25. Association for Computing Machinery, New York, NY, USA, 2025. doi: 10.1145/3706598.3713777 9

[37] L. Sidenmark and H. Gellersen. Eye, Head and Torso Coordination During Gaze Shifts in Virtual Reality. *ACM Trans. Comput.-Hum. Interact.*, 27(1), Dec. 2019. doi: 10.1145/3361218 9

[38] A. Steed. Towards a General Model for Selection in Virtual Environments. In *3D User Interfaces (3DUI'06)*, pp. 103–110, 2006. doi: 10.1109/VR.2006.134 1

[39] F. Steinicke, T. Ropinski, and K. Hinrichs. *Object Selection in Virtual Environments Using An Improved Virtual Pointer Metaphor*, pp. 320–326. Springer Netherlands, Dordrecht, 2006. doi: 10.1007/1-4020-4179-9_46 2

[40] T. Wan, L. Zhang, H. Yang, P. Irani, L. Yu, and H.-N. Liang. Exploration of Foot-based Text Entry Techniques for Virtual Reality Environments. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24. Association for Computing Machinery, New York, NY, USA, 2024. doi: 10.1145/3613904.3642757 9

[41] D. Wang, L. Qiu, B. Li, Q. Liu, X. Wei, J. Chen, Z. Wang, and M. Fan. FocalSelect: Improving Occluded Objects Acquisition with Heuristic Selection and Disambiguation in Virtual Reality . *IEEE Transactions on Visualization & Computer Graphics*, (01):1–11, Mar. 2025. doi: 10.1109/TVCG.2025.3549554 1, 2, 9

[42] D. Wolf, J. Gugenheimer, M. Combosch, and E. Rukzio. Understanding the Heisenberg Effect of Spatial Interaction: A Selection Induced Error for Spatially Tracked Input Devices. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, p. 1–10. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3313831.3376876 1, 2, 3, 6, 7

[43] D. Yu, T. Dingler, E. Velloso, and J. Goncalves. Object Selection and Manipulation in VR Headsets: Research Challenges, Solutions, and Success Measurements. *ACM Comput. Surv.*, 57(4), Dec. 2024. doi: 10.1145/3706417 2

[44] D. Yu, J. Roberts, K. Hornbæk, and J. Bergström. Deriving Selection Techniques for GUIs based on the Multiple Process Model. CHI '25. Association for Computing Machinery, New York, NY, USA, 2025. doi: 10.1145/3706598.3713089 7

[45] D. Yu, Q. Zhou, J. Newn, T. Dingler, E. Velloso, and J. Goncalves. Fully-Occluded Target Selection in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 26(12):3402–3413, 2020. doi: 10.1109/TVCG.2020.3023606 1

[46] B. Zhang, M. Sun, Z. Li, Y. Jin, and Y. Tian. GazeBubble: Exploring Bubble Mechanism for 3D Gaze Pointing to Enhance Target Acquisition in Virtual Reality. *International Journal of Human-Computer Interaction*, pp. 1–32, 07 2025. doi: 10.1080/10447318.2025.2526594 1, 2

[47] T. Zhang, Z. Hu, A. Gupta, C.-H. Wu, H. Benko, and T. R. Jonker. RIDS: Implicit Detection of a Selection Gesture Using Hand Motion Dynamics During Freehand Pointing in Virtual Reality. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, UIST '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3526113.3545701 9

[48] F. Zhu, L. Sidenmark, M. Sousa, and T. Grossman. PinchLens: Applying Spatial Magnification and Adaptive Control-Display Gain for Precise Selection in Virtual Reality. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 1221–1230, 2023. doi: 10.1109/ISMAR59233.2023.00139 1