This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2025.3622042

Reevaluating the Gaze Cursor in Virtual Reality: A Comparative Analysis of Cursor Visibility, Confirmation Mechanisms, and Task Paradigms

Yushi Wei, Rongkai Shi, Sen Zhang, Anil Ufuk Batmaz, Pan Hui, and Hai-Ning Liang, *Member, IEEE*

Abstract—Cursors and how they are presented significantly influence user experience in both VR and non-VR environments by shaping how users interact with and perceive interfaces. In traditional interfaces, cursors serve as a fundamental component for translating human movement into digital interactions, enhancing interaction accuracy, efficiency, and experience. The design and visibility of cursors can affect users' ability to locate interactive elements and understand system feedback. In VR, cursor manipulation is more complex than in non-VR environments, as it can be controlled through hand, head, and gaze movements. With the arrival of the Apple Vision Pro, the use of gaze-controlled non-visible cursors has gained some prominence. However, there has been limited exploration of the effect of this type of cursor. This work presents a comprehensive study of the effects of cursor visibility (visible vs. invisible) in gaze-based interactions within VR environments. Through two user studies, we investigate how cursor visibility impacts user performance and experience across different confirmation mechanisms and tasks. The first study focuses on selection tasks, examining the influence of target width, movement amplitude, and three common confirmation methods (air tap, blinking, and dwell). The second study explores pursuit tasks, analyzing cursor effects under varying movement speeds. Our findings reveal that cursor visibility significantly affects both objective performance metrics and subjective user preferences, but these effects vary depending on the confirmation mechanism used and task type. We propose eight design implications based on our empirical results to guide the future development of gaze-based interfaces in VR. These insights highlight the importance of tailoring cursor metaphors to specific interaction tasks and provide practical guidance for researchers and developers in optimizing VR user interfaces.

Index Terms—Virtual Reality, Eye Tracking, Cursor Visibility, Gaze Selection, Target Acquisition, Head-mounted Display

1 Introduction

As part of the "windows, icons, menus, pointers" paradigm, a cursor (or pointer) has always been one of the most fundamental components in graphical user interfaces [53]. By serving as a visual indicator of user attention and facilitating precise input, cursors enhance both interaction accuracy and efficiency [20,37]. When it comes to immersive virtual reality (VR) environments, cursors are no longer provided in the conventional shape of a mouse pointer but as the pointing end of the raycasting metaphor. Still, it helps users know where they are aiming and need to focus, and allows them to adjust their pointing directions.

As VR and similar immersive technologies evolve, additional pointing approaches using human body inputs, such as head and eye gaze, are becoming viable alternatives [46,55,58,59]. Among these, eye gaze stands out due to its distinct advantages. Unlike hand- or head-based pointing, which typically requires a visible cursor to compensate for the lack of direct spatial awareness, such as mapping hand movement captured with a mouse to a cursor on screen, eye gaze is inherently aligned with visual attention, allowing users to naturally know where they are looking [22,59]. This intuitive coupling could ideally eliminate the need for intermediate mappings [10,48]. However, human eyes naturally perform saccades, rapid, jump-like motions, along with continuous micro-movements [22]. While this dynamic behavior can enhance interaction speed, it may also compromise accuracy and increase confusion due to instability during the interaction process [16,43]. Contrary to the ideal assumption that gaze input requires no cursor, the instability caused by saccades and micro-movements may make a visible cursor beneficial in practice, by serving as a perceptual reference and facilitating user adaptation to gaze-related noise.

Taken together, the usefulness of a visible cursor in gaze-based interactions has been debated, and whether to visualize the cursor remains inconsistent. In academic research, several studies explicitly reference whether a cursor is visible [9,17,29], while others do not mention cursor visibility at all [8,47]. This difference would undermine the generalizability of research findings. On the other hand, in industry practice, some VR headsets show a visible cursor (e.g., Pimax), whereas others default to none (e.g., Apple Vision Pro). Such variance may affect the consistency of user experience across diverse devices. Although some recent work has examined the effect of cursor visibility in gaze-based interaction, not just in desktop systems [45] but also in VR systems [2], we found two main gaps in their findings: (1) a lack of discussion in confirmation mechanisms, (2) a lack of discussion in other types of tasks. These gaps underscore the need for more systematic investigations into cursor visibility, thereby paving the way for establishing best practices in gaze-based interaction for VR systems.

In this work, we aim to bridge existing gaps and offer practical design guidance for researchers and developers. To this end, we first reviewed prior work on gaze-based interaction (Sec. 2) to identify its fundamental tasks-selection and pursuit-which constitute the core of most interaction contexts. Building on these insights, we conducted two user studies to examine how cursor visibility affects user performance and experience. The first study focused on a selection task involving different interface conditions, such as target width and movement amplitude, and three confirmation mechanisms: air tap, blinking, and dwell (Sec. 3). The second study investigated a pursuit task that varied in pursuit speed (Sec. 4). Our findings from both studies revealed that cursor visibility can significantly affect both objective and subjective results, but these effects vary depending on the confirmation mechanism and task. Based on these empirical insights, eight design implications were proposed to guide the design and development of future gaze-based interfaces (Sec. 5).

The contributions of this paper include:

- Deeper insights and empirical evidence on the confirmation mech-

- *Yushi Wei, Rongkai Shi, Sen Zhang, Pan Hui, and Hai-Ning Liang are with the Computational Media and Arts Thrust in the Information Hub at The Hong Kong University of Science and Technology (Guangzhou), China.*
- *Sen Zhang is also with the School of Advanced Technology at Xi'an Jiaotong-Liverpool University, China.*
- *Anil Ufuk Batmaz is with the Department of Computer Science & Software Engineering at Concordia University, Canada.*
- *Corresponding authors: Hai-Ning Liang (hainingliang@hkust-gz.edu.cn).*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2025.3622042

anisms of both visible and invisible cursors in gaze-based selection tasks.

- Empirical findings on the effect of the cursor visibility in a pursuit task.
- Eight design implications for providing a reference for designers and researchers in effectively leveraging cursor metaphors in future VR systems.

2 Related Work

In this section, we discuss and synthesize related work on gaze-based interaction and cursor visibility.

2.1 Gaze-based Interaction

Recent advances in eye-tracking technology have established eye gaze as an effective input modality for user interaction in virtual environments (VEs). Numerous studies have empirically demonstrated that utilizing gaze as the input modality can reduce physical exertion while maintaining interactivity [25], offer intuitive controls that improve user experience [11], and allow faster and more efficient interactions [52]. Meanwhile, modern eye-tracking modules are commonly lightweight and minimally intrusive to be integrated into VR headsets [9]. This seamless integration helps avoid ergonomic drawbacks, spreading the adoption of gaze interaction across diverse VR tasks and applications, such as target selection [55, 59], steering [22], text editing [33], and gaming [35].

The human eyes execute rapid and discrete movements, a characteristic known as "saccadic movements" [44]. These swift, jump-like motions enable the eyes to shift focus quickly, making faster movements than by hands or head [18, 47]. However, gaze-based interaction tends to face accuracy challenges due to the inherent noise of natural eye movement, such as tremors, drifts, and microsaccades [13, 43]. These constant micro-movements generate uncertainty, reducing the overall interaction accuracy [16, 22]. Related to this is the "Midas Touch" problem, where brief, unintended gaze shifts are misinterpreted as intentional inputs, which is prominent in gaze-based interaction [54]. To better overcome these issues, researchers often break down the whole interaction flow into more controlled, lower-level tasks to examine them in isolation. Two representative tasks used for this purpose are the discrete-type *pointing selection* [61] and the continuous-type *pursuit* [21].

In pointing selection tasks, users direct their gaze toward the target and engage a specific mechanism to confirm the selection [42]. Currently, a variety of confirmation methods have been proposed. The most prevalent confirmation methods include *air tap* [28, 34, 38], *blinking* [14, 30, 50, 60], and *dwell* [23, 29, 36, 40], which have not just been proposed in academia but have also been deployed in commercial devices (e.g., Apple Vision Pro[1]). Specifically, air tap allows users to confirm their action by pinching their index finger and thumb [34], while blink demands briefly closing their eyes after targeting [55], and dwell requires maintaining their gaze on a target for a set duration [40]. Existing studies conducted with a visible cursor consistently show that the three confirmation mechanisms differ in distinct ways. Specifically, air tap often excels in selection accuracy, while exclusively eye-based confirmations like blinking and dwelling can inadvertently introduce additional gaze shifts or cause eye fatigue, especially in cluttered or complex VEs [7, 12, 55].

Pursuit is another common gaze-based interaction task in which users follow a moving stimulus with their eyes [5, 6]. Such a continuous trajectory-tracing task exists in various applications, including text entry [32], eye tracker calibration [41], viewport control [27], and user interface manipulation [15]. Compared to other gaze-based interactions, pursuit often offers higher accuracy because it leverages smooth eye movements triggered specifically by moving stimuli [15, 39]. In addition, its deliberate, continuous movements help to mitigate the "Midas Touch" problem [54]. Nevertheless, pursuit-based interaction also presents certain challenges, such as eye fatigue resulting from prolonged tracking of moving objects and potential confusion during user interactions with the under-designed interface [15, 39].

2.2 The Cursor Visibility

Cursors, also referred to as pointers, are used in nearly all interactive systems and serve to indicate the location where users intend to interact. Researchers have long refined or redesigned the cursor's metaphors to improve interaction efficiency and user experience (e.g., [20, 37]). However, there are differing opinions on whether a cursor should be displayed in gaze-based interaction, a factor known as cursor visibility, the presence or absence of a visible cursor. Since the location where users look at is exactly where users intend to interact, making the cursor not as essential as in interactions made by other inputs.

Some studies choose to display the gaze cursor [9, 29, 55], others do not [17, 24], and still others make no mention of whether a cursor was visible in their user studies [8, 47]. This lack of consensus also appears in commercial VR systems. For example, Apple Vision Pro keeps the gaze cursor invisible by default.[2] Similarly, the HTC VIVE Wave XR SDK explicitly recommends against using cursor guidance for eye tracking directly.[3] In contrast, a Pimax tutorial blog suggests using an eye cursor to remind users where they are looking, particularly in gaming contexts.[4] Meanwhile, the more common situation is that manufacturers like Pico and Meta generally leave the decision about the visibility of the cursor to designers and developers. Thus, whether to show the cursor in gaze-based interaction requires exploration and validation.

Rajashekhar et al. [45] found that hiding the cursor yielded superior performance and user experience in 2D selection tasks on a PC monitor, suggesting that the visible cursor might distract rather than assist in gaze-based selection. However, these findings may not extend to VR systems, where interactions differ substantially from PCs and other traditional interactive systems. This gap was recently filled by Fernandes et al. [2], who investigated the effect of cursor visibility on targeting and selecting virtual elements in VR. Fernandes et al. observed that showing a cursor led to the lowest performance and preference ratings among four selection feedback techniques, attributing this to a "cursor-chase effect" in which participants tried to follow the moving cursor rather than focusing naturally on targets [2]. However, their study has a few limitations. First, the visible cursor condition provided only the cursor itself. This disparity indicates an inconsistent study design and is incongruent with real-world systems, which typically provide additional feedback cues when a visible cursor is employed. Meanwhile, the study focused on a single selection task but overlooked other potential interaction tasks. Other works have highlighted the potential benefits of displaying the cursor, suggesting that the effectiveness of cursor visibility may be task-specific [3, 49]. Finally, only one confirmation mechanism was examined, despite earlier research indicating that different confirmation modalities can significantly affect task outcomes.

2.3 Synthesis of the Literature

In the above literature review, we first distilled two fundamental gaze-based tasks (pointing selection and pursuit) and three widely adopted confirmation mechanisms (air tap, blinking, and dwell). Then, we observed how cursor metaphors influence users' performance and behavior, highlighting the ongoing debate on whether a gaze cursor should be visible. Recent studies have begun to investigate cursor visibility in VR-based gaze interaction; however, we find that (1) the current empirical findings of visualizing the cursor in gaze-based pointing selection lack rigorous investigations using different confirmation mechanisms and with consistent supplementary feedback, and (2) the effect of cursor

[1] https://www.apple.com/apple-vision-pro/guided-tour/

[2] https://support.apple.com/guide/apple-vision-pro/use-a-pointer-to-navigate-tan3869c8a85/visionos

[3] https://hub.vive.com/storage/docs/en-us/UnityXR/UnityXREyeTracking.html

[4] https://pimax.com/blogs/blogs/eye-tracking-on-vr-virtual-reality-head srsltid=AfmBOopNqFGzbo8kN9qep4ASgSvHT5zKmfcOTijQs8tVpLdM5va_X1n2

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2025.3622042

visibility in gaze-based interaction in continuous tasks, such as pursuit tasks, is unknown. Thus, we aim to fulfill the above gaps in this work.

3 User Study 1: Pointing Selection

In this study, we aim to investigate the effects of cursor visibility and confirmation mechanisms on user performance and preferences in gaze-selection tasks. To do so, we conduct a Fitts' Law study with movement amplitude and target width as two task-related factors in addition to the cursor visibility and confirmation mechanisms. We address the following research questions (denoted by **RQ#**):

- **RQ1**: How does user performance in Fitts' law selection tasks differ among the three confirmation mechanisms when the cursor is visible? And when it is invisible?
- **RQ2**: Is the performance difference across confirmation mechanisms consistent when the cursor is visible and invisible?
- **RQ3**: How does cursor visibility affect user performance in Fitts' law selection tasks across different task characteristics (i.e., movement amplitude and target width) with each confirmation mechanism?
- **RQ4**: How does user experience in selection tasks vary across each combination of cursor visibility and confirmation mechanisms?

3.1 Participants and Apparatus

We recruited 28 participants through convenience sampling without specific inclusion or exclusion criteria, except for a willingness to participate. Participants' ages ranged from 18 to 34 years ($M = 22.42, SD = 2.92$). Based on the pre-experiment questionnaire, participants' ratings of their prior experience with VR HMDs on a 5-Likert scale were $M = 2.23$, $SD = 1.22$. All participants had normal vision and were able to see the objects in the VEs clearly. To minimize the risk of calibration or tracking errors, none wore glasses or contact lenses during the study.

We used a Meta Quest Pro VR HMD for this user study, which offers a horizontal field of view (FoV) of 106° and a vertical FoV of 95.57°, along with a display resolution of 1800×1920 pixels per eye. The experimental program was developed using Unity3D with the Oculus Integration SDK (version 57.0.2) and deployed on a laptop equipped with an AMD Ryzen AI 9 HX 370 processor and a dedicated NVIDIA GTX 4070 graphics card. The headset was connected to the laptop using a 3-meter USB-C cable.

3.2 Task

In this study, we employed a modified Fitts' ring design (ISO9241-411), a widely recognized standard paradigm for selection experiments, to systematically assess cursor impact [1]. A total of 11 grey virtual spheres were positioned in front of the participants. In each trial, one target sphere would turn blue, signaling that the target participants needed to select (see Fig. 1). Once participants confirmed their selection, the next target appeared opposite the just-selected one in a clockwise sequence. Unlike previous studies that used only a cursor and provided no additional feedback [2], our task incorporated feedback mechanisms to better simulate the selection process in the state-of-the-art real-world applications. Specifically, we employed boundary highlighting in both visible and invisible cursor conditions. Boundary highlighting is a representative selection feedback widely used in both academia [51,57] and industry[5], where hovering over a target triggered an outline. In our implementation, the outline was rendered in red. In addition, a short beep sound was provided once the selection was confirmed. Participants were instructed to select the target comfortably and naturally, as in their daily interactions, to reflect users' natural behavior. The trial would move to the next after participants confirmed a selection regardless of its correctness.

[5]For an industry example, see https://developer.apple.com/design/human-interface-guidelines/focus-and-selection

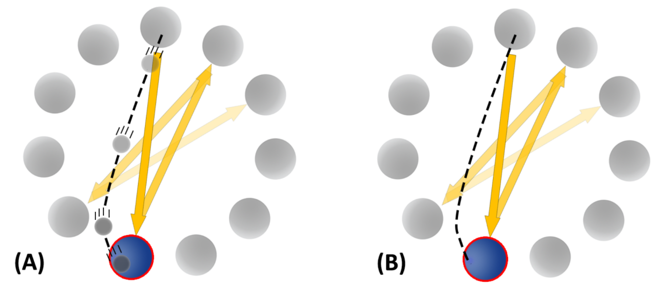


Fig. 1: Illustrations of the experimental task in Study 1. (A) The visible cursor condition. (B) The invisible cursor condition. Spherical objects are arranged in a circle. The blue sphere is the target in the current trial, and the gray spheres represent candidate targets in other trials. Yellow arrows indicate the sequence of target selections. The gray dot denotes the cursor in the visible cursor condition, and the dashed line illustrates the eye-movement path (not shown to participants). When the cursor reaches a target, that target is highlighted with a red outline.

3.3 Design

Our study used a $3 \times 3 \times 3 \times 2$ within-subject design with four independent variables: Movement Amplitude (A), Target Width (W), Confirmation Mechanisms (CM), and Cursor Visibility (CV). Next, we detail these variables, their conditions, and their rationale.

Movement Amplitude (A; 20°, 30°, and 40°) is the distance between targets in two sequential trials. We selected visual angles that incrementally increased the demand for eye-head coordination. At 20°, participants could comfortably see the entire ring with only their eyes. As amplitude increased to 30°, the ring's boundary approached the limit of natural eye movement, prompting slight head movement [26]. At 40°, this constraint became more pronounced, requiring more head movement to maintain a clear view of the target with a comfortable posture.

Target Width (W; 1.5°, 2.75°, and 4°) represents the width of each object, that is, the visual diameter of spheres. The first two conditions, 1.5°, 2.75° were chosen from the study described by Wei et al. [55] to represent the difficult and medium selection difficulties, respectively. The largest width 4° was added in this study to include a clearly easy condition.

We selected the three most commonly used Confirmation Mechanisms (CM) in eye-based selection—air tap, blinking, and dwell, as discussed in Sec. 2.1. The air tap required a pinch of the index finger and thumb to confirm a selection. The blinking was triggered when a participant's eye remained closed for 0.15s [55]. The dwell mechanism required participants to maintain their eye gaze within the target's boundary for 0.6s [29]. To provide participants with hints on the selection process of the dwell technique, the boundary color was initially lighter and gradually intensified to indicate the completion of the selection. Notably, unlike air tap and blinking, where participants may select even when the cursor is outside the target, dwell-based selection requires that the cursor remain entirely within the target boundary. To accommodate situations where dwell selection is particularly challenging (e.g., small target widths or target positions requiring greater eye–head coordination), participants are permitted to skip the current trial.

Finally, with respect to Cursor Visibility (CV; visible and invisible), we designed the visible cursor as semi-transparent gray circles with a visual size of 0.5°, ensuring sufficient visibility while minimizing occlusion [2]. In contrast, the invisible cursor condition does not show the cursor.

The order of the CM conditions was first randomized across participants. Within each CM, the appearance of CV was counterbalanced using a Latin square design. Within each $CM \times CV$ condition, we randomized the sequence of $W \times A$ combinations, each of which had 11 trials. This procedure ensured that participants could neither anticipate nor infer the order of conditions. Data from the first target was excluded because the distance between the start button and this target was only half of the intended distance. Consequently, a total of 3 W × 3 A × 3 CM × 2 CV × 10 trials × 28 participants = 15,120 trials were collected.

3.4 Procedure

Before the experiment began, participants filled out a questionnaire to collect demographic information. The experimenter then introduced the study's objectives, the tasks, and the devices being used. Participants then completed the tasks for each condition. Each $CM \times CV$ condition started with the eye calibration procedure, after which was a training session for participants to familiarize themselves with the task, the confirmation mechanism, and cursor visibility. Once they felt ready, the formal sessions commenced. During the formal sessions, if the participants reported a misalignment between their perceived and the displayed gaze positions or the experimenter found large deviations between them from the real-time data, participants were required to redo the eye calibration and the formal trials in the current condition. Participants took breaks after each condition, during which they completed questionnaires to report their subjective experiences and had enough rest. At the end of the experiment, we conducted a short interview to gain deeper insights into participants' subjective feelings regarding each condition. The entire experiment lasted approximately 45 minutes per participant.

3.5 Measurements

To systematically investigate users' selection performance and experience, we collected five objective measures and two subjective measures as our dependent variables.

Objective data were gathered through our experimental program. *Selection time* was defined as the average trial completion duration. *Success rate* reflected selection accuracy, calculated as the proportion of successful trials to the total number of trials in each condition. By successful, we meant the cursor fell within the target boundary when confirming selection. *Throughput* measures selection performance by combining speed and accuracy (please refer to our appendix for its calculation). In each condition, we also calculated *reentry times*—the average number of times the cursor reentered the target boundary. Reentry time indicates the level of effort and difficulty, with more reentries suggesting greater effort. *Endpoint deviation*, the last objective measure, was the distance (in visual degrees) between the selected location and the target's center [16]. Endpoint deviation was not recorded for dwell because it involves a selection process without a discrete trigger event, or, say, it does not generate a clear endpoint.

Subjective data were collected via questionnaires. We used the Raw NASA-TLX questionnaire [19] to assess participants' perceived *workload*, and a questionnaire designed by Piumsomboon et al. [42] to measure different aspects of the *usability*.

3.6 Results

Before analyzing the results, we discarded outliers using the same criterion as prior work [55]. This removed 247 air tap trials (112 invisible, 135 visible; 4.4% and 5.3% of their respective totals), 267 blinking trials (146 invisible, 121 visible; 5.8% and 4.8%), and 144 dwell trials (97 invisible, 47 visible; 3.8% and 1.9%). All 34 trials that were skipped in the dwell condition occurred with an invisible cursor and were likewise excluded.

3.6.1 Results on RQ1: Mechanism Differences Varied by Visibility

To answer RQ1, we conducted one-way ANOVA tests, or Friedman tests when normality assumptions were violated, to examine the effect of confirmation mechanisms (CM) for user performance in visible cursor and invisible cursor conditions and visualized the results in Fig. 2. Results showed that significant effects of CM on all measures within two CV conditions, except the endpoint deviation in the visible cursor condition. The test results are summarized in Tab. 1. When significant effects were observed, pairwise comparisons were further performed, with all significant results reported in Tab. 2.

Table 1: Significant statistical test results: the effects of confirmation mechanisms (CM) on different measures, conducted separately for visible cursor and invisible cursor conditions.

	Dependent Variables	Statistics
Visible Cursor	Selection Time	$\chi^2_2 = 31.10,\ p < 0.001,\ W = 0.56$
	Success Rate	$\chi^2_2 = 43.70,\ p < 0.001,\ W = 0.78$
	Throughput	$F(2,54) = 11.583,\ p < 0.001,\ \eta^2_p = 0.300$
	Reentry Time	$\chi^2_2 = 39.20,\ p < 0.001,\ W = 0.70$
Invisible Cursor	Selection Time	$\chi^2_2 = 36.90,\ p < 0.001,\ W = 0.66$
	Success Rate	$\chi^2_2 = 47.80,\ p < 0.001,\ W = 0.85$
	Throughput	$\chi^2_2 = 14.90,\ p < 0.001,\ W = 0.27$
	Reentry Time	$\chi^2_2 = 39.50,\ p < 0.001,\ W = 0.71$
	Endpoint	$\chi^2_1 = 11.60,\ p < 0.001,\ W = 0.40$

Table 2: Pairwise comparisons: the effects of confirmation mechanisms (CM) on different measures, conducted separately for visible cursor and invisible cursor conditions.

		Comparisons	Statistics
Visible Cursor	Selection Time	Dwell > Air Tap	$Z = 9.25,\ p < 0.001,\ r = 0.538$
		Dwell > Blinking	$Z = 8.35,\ p < 0.001,\ r = 0.526$
	Success Rate	Dwell > Air Tap	$Z = 11.7,\ p < 0.001,\ r = 0.867$
		Dwell > Blinking	$Z = 12.3,\ p < 0.001,\ r = 0.867$
		Air Tap > Blinking	$Z = 3.17,\ p = 0.002,\ r = 0.213$
	Throughput	Dwell > Air Tap	$t_{(251)} = 5.99,\ p = 0.006,\ d_z = 0.377$
		Dwell > Blinking	$t_{(251)} = 7.81,\ p < 0.001,\ d_z = 0.492$
		Air Tap > Blinking	$t_{(251)} = 2.2,\ p = 0.028,\ d_z = 0.138$
	Reentry Time	Dwell > Air Tap	$Z = 11.1,\ p < 0.001,\ r = 0.698$
		Dwell > Blinking	$Z = 11.5,\ p < 0.001,\ r = 0.660$
Invisible Cursor	Selection Time	Dwell > Air Tap	$Z = 11.7,\ p < 0.001,\ r = 0.740$
		Dwell > Blinking	$Z = 12.1,\ p < 0.001,\ r = 0.762$
	Success Rate	Dwell > Air Tap	$Z = 13.6,\ p < 0.001,\ r = 0.857$
		Dwell > Blinking	$Z = 13.7,\ p < 0.001,\ r = 0.866$
		Air Tap > Blinking	$Z = 4.78,\ p < 0.001,\ r = 0.301$
	Throughput	Air Tap > Dwell	$Z = 2.03,\ p = 0.042,\ r = 0.128$
		Air Tap > Blinking	$Z = 6.16,\ p < 0.001,\ r = 0.388$
		Dwell > Blinking	$Z = 4.22,\ p < 0.001,\ r = 0.266$
	Reentry Time	Dwell > Air Tap	$Z = 12.0,\ p < 0.001,\ r = 0.745$
		Dwell > Blinking	$Z = 12.5,\ p < 0.001,\ r = 0.788$
	Endpoint	Blinking > Air Tap	$Z = 7.32,\ p < 0.001,\ r = 0.461$

In the visible cursor condition, dwell led to significantly longer selection time than air tap and blinking (both $p < 0.001$; Fig. 2 A). The longer selection time resulted in a higher success rate. Dwell had a significantly higher success rate than air tap and blinking (both $p < 0.001$), and air tap's success rate also surpassed blinking ($p = 0.002$), as can be seen in Fig. 2 B. A significant difference among CM was also found in throughput (Fig. 2 C). Same as in the success rate, dwell outperformed both air tap ($p = 0.006$) and blinking ($p < 0.001$), and air tap also exceeded blinking ($p < 0.028$). Regarding reentry times (Fig. 2 D), dwell further yielded the most reentries, exceeding both air tap and blinking (both $p < 0.001$). However, we found no significant differences among the mechanisms in endpoint deviation (Fig. 2 E).

With an invisible cursor, CM exhibited significant effects across all five metrics. As shown in Fig. 2 F, dwell was the slowest, with selection times longer than both air tap and blinking (both $p < 0.001$). Despite this slower performance, it achieved the highest accuracy (Fig. 2 G), outperforming both air tap and blinking (all $p < 0.001$), with air tap

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2025.3622042

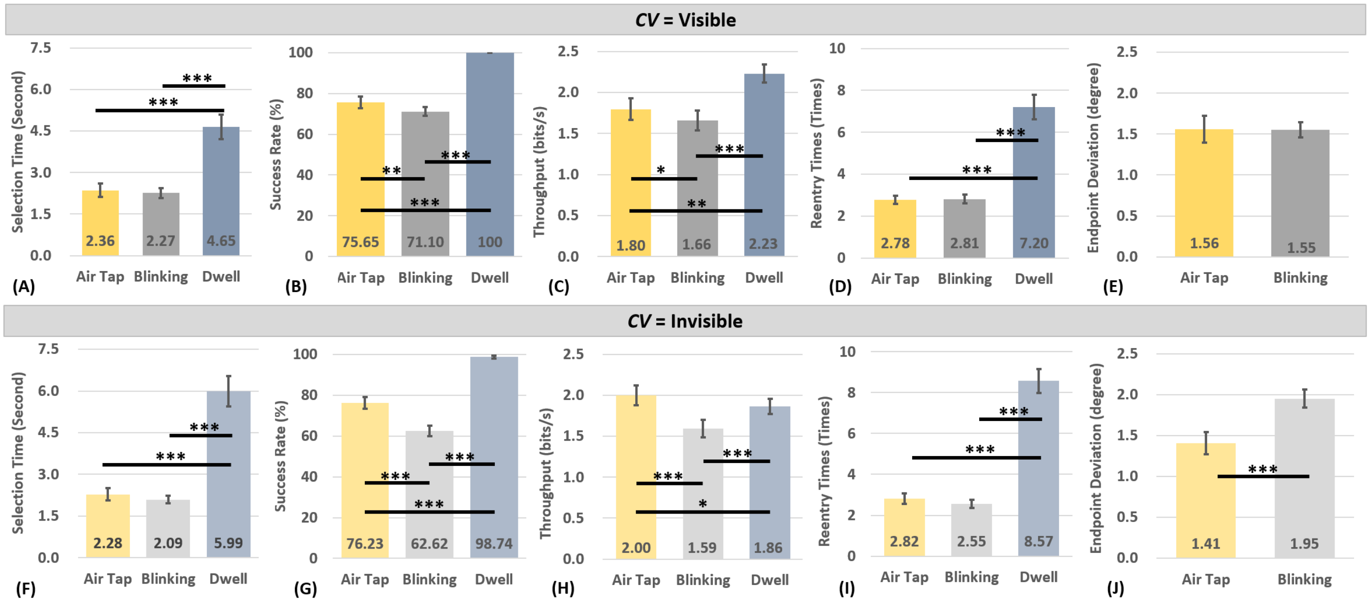


Fig. 2: Five objective metrics across three confirmation mechanisms (*CM*)
. Error bars represent the standard errors. Asterisks *, **, *** indicate significant differences at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

also surpassing blinking ($p < 0.001$). Throughput (Fig. 2 H), however, presented a different outcome: air tap produced the highest throughput, surpassing both dwell ($p = 0.042$) and blinking ($p < 0.001$), with dwell still outperforming blinking ($p < 0.001$). As for reentry times (Fig. 2 I), dwell generated the most reentries, more than air tap and blinking (both $p < 0.001$). Finally, regarding endpoint deviation (Fig. 2 J), blinking caused greater deviation than air tap ($p < 0.001$).

3.6.2 Results on RQ2: Consistency of Mechanism Differences Across Visibility Conditions

To answer RQ2, whether performance differences across confirmation mechanisms (*CM*) varied with cursor visibility (*CV*), we conducted RM-ANOVAs to test the interaction effect ($CM \times CV$). The analyses revealed significant interaction effects in three metrics: success rate ($F(2,54) = 3.87, p = .032, \eta_p^2 = .125$), throughput ($F(2,54) = 7.12, p = .002, \eta_p^2 = .209$), and endpoint deviation ($F(1,27) = 11.45, p = .002, \eta_p^2 = .298$). These results indicate that cursor visibility modulated performance differences among the three confirmation mechanisms under these three metrics.

Specifically, the success rates in both *CV* conditions showed the same trends: dwell > air tap > blinking. However, the gap between air tap and blinking was enlarged when the cursor was invisible, indicating that invisibility amplified their performance difference. The throughput of the three confirmation mechanisms changed when the cursor was visible (dwell > air tap > blinking) versus when it was invisible (air tap > dwell > blinking). Finally, for endpoint deviation, no differences were observed in the visible condition, while in the invisible condition, we found that blinking produced significantly greater deviation than air tap.

3.6.3 Results on RQ3: Visibility Outcomes Varied by Target Size and Amplitude

To answer RQ3, repeated measures ANOVAs (RM-ANOVAs) were conducted for each dependent measure within each *CM* to test the main effect of *CV* and interaction effects of $A \times CV$ and $W \times CV$. When the assumptions of RM-ANOVAs were not met, the Aligned Rank Transform (ART) procedure was applied to permit valid nonparametric analyses [56]. Pairwise comparisons were performed with Bonferroni-Holm correction, even when RM-ANOVA test results were not significant, to examine potential condition-specific differences that might be masked when data were aggregated across task parameters. We used t-tests for pairwise comparisons when the normality assumption was met and Wilcoxon signed-rank tests when it was violated. We found significant performance differences between visible and invisible cursor conditions when *CM* was blinking and dwell (see Fig. 3, Tab. 4, and Tab. 3), but no significant differences emerged across all measures when the *CM* was air tap.

For blinking, we identified a clear pattern in accuracy. Blinking with a visible cursor consistently achieved significantly higher success rates than with an invisible cursor across all movement amplitudes and target widths (see Fig. 3 B). Similarly, blinking with a visible cursor also resulted in significantly lower deviation across all conditions, except for $A = 20°$ (Fig. 3 E). Speed-wise, we did not find such a consistent pattern. The visible cursor yielded longer selection times than the invisible cursor in isolated cases ($A = 20°$: $p = 0.013$; $W = 2.75$: $p = 0.016$; and $W = 4°$: $p < 0.001$), as can be seen from Fig. 3 A. Because of this, we did not find any significant differences in throughput between using a visible cursor or not for blinking. Finally, we found blinking with a visible cursor led to more reentries when $A = 40°$ ($p = 0.004$) and $W = 1.5°$ ($p = 0.046$).

For dwell, we found no significant differences in success rate between visible and invisible cursors except when $W = 1.5°$ (visible > invisible, $p = 0.003$). Similarly, there were no significant differences in endpoint deviation. However, selection times were significantly shorter with the visible cursor across all conditions, as shown in Fig. 3 A. It eventually contributed to a significantly higher throughput for the visible cursor across all conditions (see Fig. 3 C). We also found that dwell with a visible cursor had significantly lower reentry counts across all conditions except for when $W = 1.5°$ (see Fig. 3 D).

3.6.4 Results on RQ4: Subjective Experience

The results of RM-ANOVAs on subjective experience are reported in the text. Pairwise comparisons are summarized in Tab. 5 and Tab. 6, which present the test statistics, *p*-values, and effect sizes for workload and usability. The corresponding pairwise results are visualized in Fig. 4. For clarity, A, B, and D denote air tap, blinking, and dwell, while V and I indicate the visible and invisible cursor conditions; combinations such as AV or BI therefore represent a specific confirmation mechanism under a given cursor condition.

Workload The results showed that AV produced the lowest mental demand scores. Specifically, AV was significantly lower than both BV and DV, while AI was significantly lower than both BI and DI.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2025.3622042

Table 3: Pairwise comparisons between visible and invisible cursor conditions (CV) across amplitudes (A) and widths (W), conducted separately for each confirmation mechanism (CM). Only significant results are shown. Results are ordered by amplitude (A) first and width (W) second; within each A or W condition, Blinking is presented first, followed by Dwell. A = Air Tap, B = Blinking, D = Dwell, V = Visible, and I = Invisible..

	Conditions	Comparisons	Statistics
Selection Time	$A = 20°$	BI < BV	$Z = -2.75,\ p = 0.013,\ r = -0.30$
	$A = 20°$	DI > DV	$Z = 2.30,\ p = 0.021,\ r = 0.251$
	$A = 30°$	DI > DV	$Z = 1.96,\ p = 0.050,\ r = 0.213$
	$A = 40°$	DI > DV	$Z = 2.02,\ p = 0.043,\ r = 0.220$
	$W = 2.75°$	BI < BV	$Z = -2.68,\ p = 0.016,\ r = -0.293$
	$W = 4°$	BI < BV	$Z = -3.30,\ p < 0.001,\ r = -0.36$
	$W = 1.5°$	DI > DV	$t(83) = 2.77,\ p = 0.006,\ d_z = 0.303$
	$W = 2.75°$	DI > DV	$Z = 1.88,\ p = 0.047,\ r = 0.205$
	$W = 4°$	DI > DV	$Z = 3.84,\ p < 0.001,\ r = 0.419$
Success Rate	$A = 20°$	BV > BI	$t(83) = 2.67,\ p = 0.009,\ d_z = 0.291$
	$A = 30°$	BV > BI	$t(83) = 2.79,\ p = 0.006,\ d_z = 0.304$
	$A = 40°$	BV > BI	$t(83) = 3.25,\ p = 0.001,\ d_z = 0.355$
	$W = 1.5°$	BV > BI	$t(83) = 2.55,\ p = 0.012,\ d_z = 0.278$
	$W = 2.75°$	BV > BI	$t(83) = 2.31,\ p = 0.023,\ d_z = 0.254$
	$W = 4°$	BV > BI	$t(83) = 4.26,\ p < 0.001,\ d_z = 0.465$
	$W = 1.5°$	DV > DI	$Z = 7.96,\ p = 0.003,\ r = 0.869$
Throughput	$A = 20°$	DV > DI	$t(83) = 3.93,\ p < 0.001,\ d_z = 0.429$
	$A = 30°$	DV > DI	$t(83) = 2.83,\ p = 0.005,\ d_z = 0.308$
	$A = 40°$	DV > DI	$Z = 2.85,\ p = 0.004,\ r = 0.311$
	$W = 1.5°$	DV > DI	$Z = 2.23,\ p = 0.025,\ r = 0.244$
	$W = 2.75°$	DV > DI	$t(83) = 3.01,\ p = 0.003,\ d_z = 0.329$
	$W = 4°$	DV > DI	$t(83) = 4.51,\ p < 0.001,\ d_z = 0.492$
Reentry Time	$A = 40°$	BV > BI	$Z = 2.85,\ p = 0.004,\ r = 0.311$
	$A = 20°$	DI > DV	$Z = 2.46,\ p = 0.007,\ r = 0.268$
	$A = 30°$	DI > DV	$Z = 1.60,\ p = 0.049,\ r = 0.175$
	$A = 40°$	DI > DV	$Z = 2.18,\ p = 0.003,\ r = 0.238$
	$W = 1.5°$	BV > BI	$Z = 2.14,\ p = 0.046,\ r = 0.234$
	$W = 2.75°$	DI > DV	$Z = 2.70,\ p = 0.003,\ r = 0.238$
	$W = 4°$	DI > DV	$Z = 3.51,\ p < 0.001,\ r = 0.238$
Endpoint	$A = 30°$	BI > BV	$t(83) = 2.76,\ p = 0.007,\ d_z = 0.301$
	$A = 40°$	BI > BV	$Z = 3.20,\ p = 0.001,\ r = 0.349$
	$W = 1.5°$	BI > BV	$Z = 2.34,\ p = 0.019,\ r = 0.255$
	$W = 2.75°$	BI > BV	$Z = 2.91,\ p = 0.003,\ r = 0.317$
	$W = 4°$	BI > BV	$t(83) = 2.92,\ p = 0.004,\ d_z = 0.318$

Moreover, BV and BI yielded significantly lower scores than DV and DI, respectively; meanwhile, DI exhibited the highest workload, significantly exceeding AI, BI, and DV.

Regarding physical demand, AV exhibited the lowest scores, requiring significantly less physical effort than both BV and DV. In addition, AI also demonstrated lower physical demand than DI.

Concerning temporal demand, AV produced the lowest scores, which were significantly lower than both those produced by BV and DV. In contrast, DI showed the highest temporal demand, scoring significantly higher than AI, BI, and DV.

In terms of frustration ratings, the lowest frustration was observed for AV, which was significantly lower than both BV and DV. Moreover, AI was associated with lower frustration than both BI and DI. Significant differences were also found between BI and DI, as well as between DV and DI.

For performance ratings, AV achieved the lowest ratings, which were significantly lower than those of BV and DV. AI received the second-lowest ratings, with ratings significantly lower than those for BI and DI. A significant difference was also identified between DV and DI.

For effort demand, AV required the least effort, with scores significantly lower than those of DV. Moreover, AI ranked second, with significantly lower effort demand than both BI and DI. In contrast, DI required the highest effort, scoring significantly higher than AI, BI, and DV.

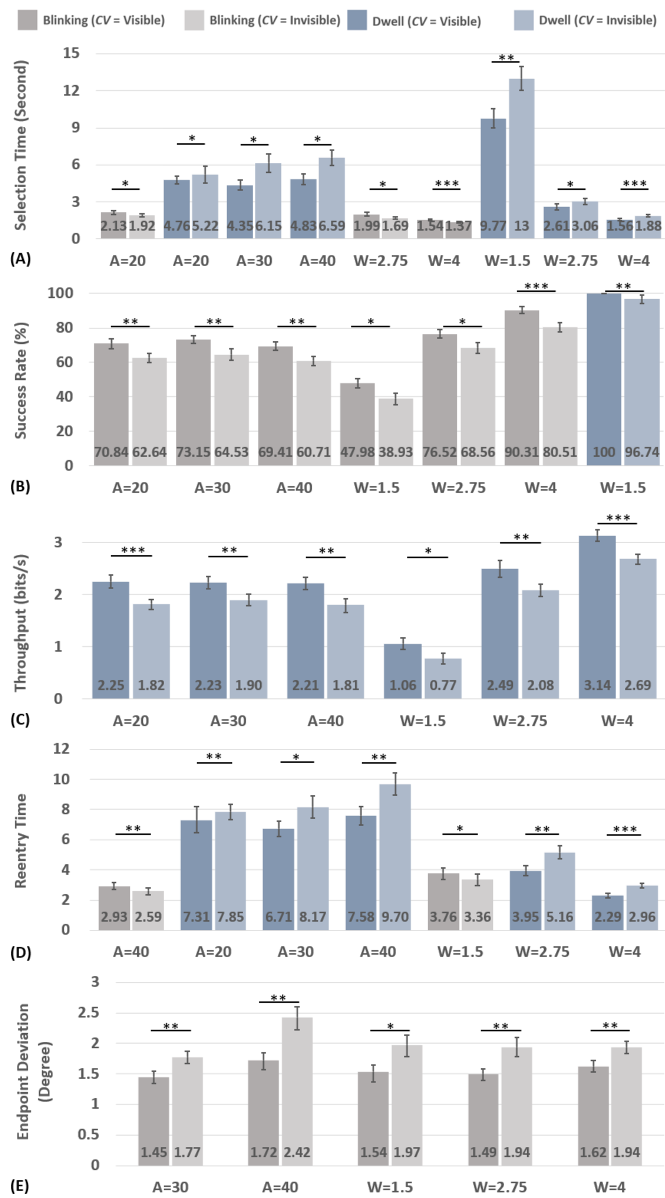


Fig. 3: Five objective metrics across two cursor conditions. Error bars represent the standard error. Asterisks (*, **, ***) indicate significant differences at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. Only significant results are reported, ordered by amplitude (A) and then by width (W). Within each A or W condition, blinking results precede dwell.

DV.

The overall scores revealed that AV yielded significantly lower scores than both BV and DV. Similarly, AI produced lower scores than both BI and DI. A significant difference was also identified between BV and DV. Conversely, DI received the highest overall scores, exceeding AI, BI, and DV.

Usability Concerning Q1 "It feels natural to use", the analysis revealed that DI received the lowest naturalness rating, being significantly lower than both AI and BI. In contrast, AV was rated as the most natural method, significantly outperforming both BV and DV.

With respect to Q2 "I could interact precisely", among all conditions, only DI demonstrated a significant drop in interaction precision, scoring lower than both AI and DV.

In terms of Q3 "It was easy for me to use", AV was rated significantly higher than both BV and DV. Similarly, AI outperformed both BI and

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2025.3622042

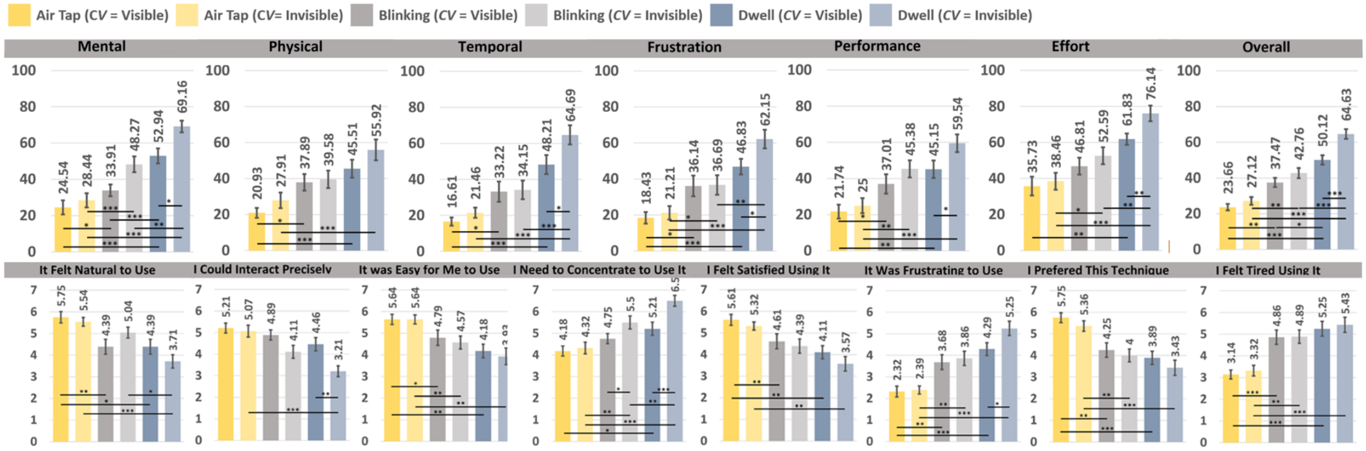


Fig. 4: Subjective results. Error bars represent the standard error. Asterisks (*, **, ***) indicate significant differences at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

Table 4: Significant RM-ANOVAs results for CV, $A \times CV$, and $W \times CV$ across CM conditions in Study 1 ($\alpha = .05$).

	Measure	Effect	F-statistics	p	η_p^2
Blinking	Selection Time	$W \times CV$	$F(2,54) = 3.289$	0.047	0.108
	Success Rate	CV	$F(2,54) = 4.32$	0.032	0.138
	Endpoint Deviation	CV $A \times CV$	$F(1,27) = 11.972$ $F(2,54) = 4.194$	0.002 0.020	0.307 0.134
Dwell	Selection Time	CV $A \times CV$	$F(1,27) = 4.32$ $F(2,54) = 3.565$	0.032 0.035	0.138 0.117
	Throughput	CV	$F(1,27) = 8.715$	0.006	0.244
	Reentry Time	CV $W \times CV$	$F(1,27) = 8.647$ $F(2,54) = 4.608$	0.007 0.014	0.243 0.146

DI, with higher ease-of-use ratings.

With respect to Q4 "I need to concentrate to use it", DI was rated the highest, significantly exceeding DV, BI, and AI. Additionally, DV also showed significantly higher scores than AV, while BI was rated significantly above both BV and AI.

For Q5 "I felt satisfied using it", AV was rated significantly higher than both BV and DV. Additionally, AI scored significantly higher than DI.

Regarding Q6 "It was frustrating to use", AV received the lowest frustration ratings, being significantly lower than both BV and DV. Likewise, AI was associated with significantly less frustration than both BI and DI. Notably, DV also outperformed DI, with lower frustration levels.

Concerning Q7 "I preferred this technique", a higher preference was expressed for AV compared to both BV and DV. A similar trend was observed under the invisible condition, where AI was significantly preferred over both BI and DI.

In terms of Q8 "I felt tired using it", AV resulted in the lowest tiredness ratings, being significantly lower than both BV and DV. Similarly, AI was associated with significantly less fatigue than both BI and DI, suggesting that the air tap mechanism consistently reduced perceived exhaustion across both visibility conditions.

3.7 Discussion

3.7.1 Discussion of RQ1 and RQ2: Differences in Mechanism Performance Between Visible and Invisible Cursors

Our results in Sec. 3.6.1 and Sec. 3.6.2 provide answers to RQ1 and RQ2 (Sec. 3). For RQ1, we found significant differences among the confirmation mechanisms across all metrics, except for endpoint devia-

Table 5: Pairwise comparisons of workload ratings across CM and CV. A = Air Tap, B = Blinking, D = Dwell, V = Visible, and I = Invisible.

	Comparisons	Statistics
Mental	AV < BV	$t(27) = -2.70,\ p = 0.047,\ d_z = -0.509$
	AV < DV	$t(27) = -5.92,\ p < 0.001,\ d_z = -1.12$
	AI < BI	$t(27) = -4.70,\ p < 0.001,\ d_z = -0.887$
	AI < DI	$t(27) = -7.48,\ p < 0.001,\ d_z = -1.41$
	BV < DV	$t(27) = -5.03,\ p < 0.001,\ d_z = -0.950$
	BI < DI	$t(27) = -3.92,\ p = 0.003,\ d_z = -0.741$
	DI > DV	$t(27) = 3.35,\ p = 0.012,\ d_z = 0.632$
Physical	AV < BV	$t(27) = -3.49,\ p = 0.018,\ d_z = -0.659$
	AV < DV	$t(27) = -5.29,\ p < 0.001,\ d_z = -1.00$
	AI < DI	$t(27) = -4.87,\ p < 0.001,\ d_z = -0.921$
Temporal	AV < BV	$Z = -2.73,\ p = 0.044,\ r = -0.516$
	AV < DV	$t(27) = -6.13,\ p < 0.001,\ d_z = -1.16$
	DI > AI	$t(27) = 8.69,\ p < 0.001,\ d_z = 1.64$
	DI > BI	$t(27) = 5.37,\ p = 0.006,\ d_z = 1.02$
	DI > DV	$Z = 2.91,\ p = 0.029,\ r = 0.549$
Frustration	AV < BV	$t(27) = -3.31,\ p = 0.021,\ d_z = -0.625$
	AV < DV	$t(27) = -5.89,\ p < 0.001,\ d_z = -1.11$
	AI < BI	$t(27) = -2.88,\ p = 0.045,\ d_z = -0.545$
	AI < DI	$t(27) = -7.95,\ p < 0.001,\ d_z = -1.50$
	BI < DI	$t(27) = -4.00,\ p = 0.004,\ d_z = -0.756$
	DV < DI	$t(27) = -3.22,\ p = 0.023,\ d_z = -0.609$
Performance	AV < BV	$t(27) = -3.21,\ p = 0.027,\ d_z = -0.606$
	AV < DV	$t(27) = -4.51,\ p = 0.001,\ d_z = -0.853$
	AI < BI	$Z = -3.59,\ p = 0.003,\ r = -0.679$
	AI < DI	$t(27) = -6.49,\ p < 0.001,\ d_z = -1.23$
	DV < DI	$t(27) = -3.31,\ p = 0.023,\ d_z = -0.626$
Effort	AV < DV	$t(27) = -4.46,\ p = 0.001,\ d_z = -0.842$
	AI < BI	$t(27) = -3.18,\ p = 0.025,\ d_z = -0.600$
	AI < DI	$t(27) = -6.23,\ p < 0.001,\ d_z = -1.18$
	DI > BI	$t(27) = 3.99,\ p = 0.004,\ d_z = 0.754$
	DI > DV	$Z = 3.65,\ p = 0.002,\ r = 0.689$
Overall	AV < BV	$t(27) = -3.75,\ p = 0.005,\ d_z = -0.708$
	AV < DV	$t(27) = -7.06,\ p < 0.001,\ d_z = -1.33$
	AI < BI	$t(27) = -4.10,\ p = 0.002,\ d_z = -0.775$
	AI < DI	$t(27) = -8.71,\ p < 0.001,\ d_z = -1.65$
	BV < DV	$Z = -3.02,\ p = 0.012,\ r = -0.570$
	DI > BI	$Z = 4.84,\ p < 0.001,\ r = 0.916$
	DI > DV	$Z = 4.01,\ p < 0.001,\ r = 0.757$

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2025.3622042

Table 6: Pairwise comparisons of usability ratings (Q1–Q8) across confirmation mechanisms and cursor visibility conditions. A = Air Tap, B = Blinking, D = Dwell; V = Visible, I = Invisible.

	Comparisons	Statistics
Q1 Naturalness	DI < AI	$t(27) = -5.11,\ p < 0.001,\ d_z = -0.965$
	DI < BI	$t(27) = -3.67,\ p = 0.012,\ d_z = -0.693$
	AV > BV	$Z = 3.40,\ p = 0.008,\ r = 0.642$
	AV > DV	$t(27) = 3.24,\ p = 0.034,\ d_z = 0.613$
Q2 Precision	DI < AI	$t(27) = -6.04,\ p < 0.001,\ d_z = -1.14$
	DI < DV	$Z = -3.51,\ p = 0.005,\ r = -0.664$
Q3 Ease	AV > BV	$Z = 2.85,\ p = 0.043,\ r = 0.539$
	AV > DV	$Z = 3.70,\ p = 0.003,\ r = 0.700$
	AI > BI	$Z = 3.57,\ p = 0.005,\ r = 0.674$
	AI > DI	$Z = 3.50,\ p = 0.006,\ r = 0.661$
Q4 Concentration	DI > DV	$Z = 4.08,\ p < 0.001,\ r = 0.789$
	DI > BI	$Z = 3.58,\ p = 0.003,\ r = 0.677$
	DI > AI	$t(27) = 2.64,\ p = 0.014,\ d_z = 0.498$
	DV > AV	$Z = 2.84,\ p = 0.035,\ r = 0.537$
	BI > BV	$t(27) = 3.07,\ p = 0.035,\ d_z = 0.579$
	BI > AI	$Z = 3.30,\ p = 0.008,\ r = 0.623$
Q5 Satisfaction	AV > BV	$t(27) = 3.89,\ p = 0.007,\ d_z = 0.735$
	AV > DV	$Z = 3.48,\ p = 0.006,\ r = 0.658$
	AI > DI	$t(27) = 4.02,\ p = 0.005,\ d_z = 0.760$
Q6 Frustration	AV < BV	$t(27) = -3.76,\ p = 0.008,\ d_z = -0.711$
	AV < DV	$t(27) = -5.43,\ p < 0.001,\ d_z = -1.03$
	AI < BI	$Z = -3.38,\ p = 0.008,\ r = -0.638$
	AI < DI	$Z = -4.08,\ p < 0.001,\ r = -0.772$
	DV < DI	$Z = -2.83,\ p = 0.032,\ r = -0.534$
Q7 Preference	AV > BV	$t(27) = 4.02,\ p = 0.004,\ d_z = 0.761$
	AV > DV	$t(27) = 5.52,\ p < 0.001,\ d_z = 1.04$
	AI > BI	$t(27) = 3.88,\ p = 0.005,\ d_z = 0.734$
	AI > DI	$t(27) = 4.80,\ p < 0.001,\ d_z = 0.908$
Q8 Fatigue	AV < BV	$Z = -4.13,\ p < 0.001,\ r = -0.780$
	AV < DV	$Z = -3.93,\ p < 0.001,\ r = -0.743$
	AI < BI	$Z = -3.65,\ p = 0.002,\ r = -0.690$
	AI < DI	$t(27) = -5.56,\ p < 0.001,\ d_z = -1.05$

tion in the visible condition. For RQ2, the relative performance of the three confirmation mechanisms varied between the visible and invisible cursor conditions in terms of success rate, throughput, and endpoint deviation. In contrast, no such differences were observed for selection time or reentry counts.

With or without the cursor, the three confirmation mechanisms have similar patterns in selection time: dwell required the longest durations, whereas air tap and blinking yielded similar performance. This aligns with prior work [12, 38], which attributes the prolonged dwell time to its requirement for sustained fixation to trigger a selection. In contrast, air tap and blinking terminate the trial immediately upon activation, regardless of whether the target is successfully selected.

In terms of success rate, cursor visibility led to notable performance differences across mechanisms. Dwell consistently achieved significantly higher success rates than both air tap and blinking across visible and invisible cursor conditions (all $p < 0.001$). However, the difference between air tap and blinking varied with cursor visibility, with effect sizes indicating a shift in their relative performance. Specifically, with a visible cursor, air tap outperformed blinking; this difference became more pronounced under the invisible cursor condition (see Fig. 2 B and G). We attribute this change to the decreased success rate of blinking when used with the invisible cursor. As discussed in prior work [55], compared to air tap, blinking is a subtle and often involuntary behavior that is more susceptible to external disturbances. In our context, the lack of cursor visibility may introduce greater uncertainty during selection, which likely contributes to the reduced performance observed for blinking.

Furthermore, a more pronounced difference was observed in throughput. Under the visible cursor condition, dwell yielded the highest throughput, significantly outperforming both air tap and blinking. However, air tap achieved the highest throughput, significantly exceeding that of dwell and blinking in the invisible cursor condition. Based on the definition of throughput, we attribute this shift to the increased selection time of dwell when paired with the invisible cursor, as further evidenced by Fig. 3 A, which shows significantly longer selection times under the invisible condition.

Regarding reentry times, a consistent pattern was observed across both visible and invisible cursor conditions. Dwell resulted in the highest number of reentries, whereas no significant difference was found between air tap and blinking. These higher reentry counts can be attributed to the nature of the dwell mechanism, which requires participants to repeatedly attempt selection until the target is successfully activated [38, 39].

Finally, endpoint deviation exhibited a shift in relative performance. With the cursor visible, deviations were comparable across all confirmation mechanisms. Once the cursor was hidden, however, blinking produced markedly larger endpoint deviations than air tap. As in the success rate analysis, we attribute this decline to the absence of cursor visibility, which may impair users' selection accuracy due to the inherently unstable nature of the blinking compared to air tap.

In summary, cursor visibility re-orders the relative performance of the three confirmation mechanisms rather than uniformly amplifying or diminishing them. With the cursor visible, dwell is the most reliable (highest success rate, best throughput despite the longest selection time), air tap remains competitive, and blinking lags behind. Once the cursor is hidden, the hierarchy flips, elevating air tap to the top, maintaining high success and achieving the best throughput, whereas dwell exhibits a decline in throughput, and blinking shows a reduced accuracy. In that case, we conclude that cursor visibility plays a key role in reshaping the relative ranking of confirmation mechanisms, leading to different performance outcomes under visible and invisible conditions.

3.7.2 Discussion of RQ3: Cursor Visibility Outcomes Diverge by Mechanism, Modulated by Amplitude and Width

The results in Sec. 3.6.3 provide answers to RQ3 (Sec. 3), showing that each confirmation mechanism exhibited distinct visibility-dependent patterns, most of which further varied with movement amplitude and target width.

For the air tap, no significant differences were found between visible and invisible cursors across all conditions and measures. These findings contrast with previous works [2, 45], which reported noticeably worse performance when using a visible cursor. We attribute this discrepancy to differences in feedback design: in earlier research, the visible cursor simply appeared without additional cues, whereas the invisible cursor condition provided explicit feedback (e.g., boundary highlighting). This raises a key issue: real-world applications typically implement custom feedback for visible cursors. Therefore, when using air tap as the confirmation mechanism with an explicit feedback mechanism, the cursor visibility does not appear to meaningfully affect users' performance.

Under the blinking condition, cursor visibility exerted a significant influence on performance. Compared to the invisible cursor, a visible cursor prolonged selection times in trials with lower eye–head coordination demand (small amplitude, $A = 20°$) or lower selection difficulty (large width, $W = 2.75°$ or $4°$), while increasing re-entry counts in trials posing greater task challenge ($A = 40°$, $W = 1.5°$). Nonetheless, visible cursors uniformly enhanced success rates and reduced endpoint deviations across all condition levels, indicating improved accuracy. Two factors may explain this pattern. First, with a visible cursor, participants engaged in fine-grained aiming, which extended selection times while enhancing precision, evidenced by higher success rates and smaller endpoint deviations. Second, when targets were challenging to acquire, the invisible cursor prompted more cautious strategies—characterized by increased effort and deliberate selection behaviors (reported by P2, P7, P12, P15, P22, and P28)—thereby lengthening selection times and narrowing the performance gap compared to the visible cursor.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2025.3622042

This interpretation is further supported by lower re-entry counts for the invisible cursor at $A = 40^\circ$ and $W = 1.5^\circ$.

For the dwell mechanism, the visible cursor conferred advantages across nearly all metrics, including shorter selection times, higher success rates, greater throughput, and fewer re-entries. These benefits stem from the explicit visual cues provided by the cursor, which help participants maintain fixation within the target area to trigger selection. Among the metrics, an interesting pattern occurred in the success rate, where a significant difference between the visible and invisible cursor was only observed at $W = 1.5^\circ$. We attribute the finding to the high task difficulty, which resulted in several skipped trials, exclusively at $W = 1.5^\circ$, and only under the invisible cursor condition. These skipped trials contributed to the observed success rate difference at that width. In contrast, re-entry counts at that width did not differ significantly between visibility conditions, as the high difficulty level forced participants to make multiple entry attempts regardless of cursor presence.

Overall, our findings answered RQ3. Air tap remained unaffected by cursor visibility, with metrics stable across all conditions. Blinking exhibited a task challenge-scaled visibility effect: with a visible cursor, selection time was prolonged in easier trials (characterized by small amplitudes or large widths) and re-entry counts increased in more challenging ones (large amplitudes, narrow widths), while accuracy consistently improved across all condition levels. Dwell uniformly benefited from a visible cursor. These results demonstrate that visibility effects are mechanism-specific across all confirmation methods. However, only some mechanisms, such as blinking and dwell, are further modulated by task settings, with air tap as the sole exception.

3.7.3 Discussion of RQ4: Preferences for Cursor Visibility Depend on Confirmation Mechanism

The findings in Sec. 3.6.4 answered RQ4 (Sec. 3), revealing significant differences in workload and usability across confirmation mechanisms under varying cursor visibility conditions.

Across all NASA-TLX dimensions and usability measures, the air tap paired with a visible cursor (*AV*) condition imposed the lowest mental, physical, temporal and effort demands, yielded the highest perceived performance and lowest frustration, and received the strongest overall preference. Removing the cursor (*AI*) reduced, but did not eliminate, air tap's advantage. Consistent with our objective findings, no significant differences were observed between *AV* and *AI* across all subjective measurements, indicating that cursor visibility is not a key factor when employing air tap.

Blinking consistently occupied an intermediate position within each visibility condition: its workload and usability scores were lower than air tap but higher than dwell. In support of H3, the only significant difference between visible (*BV*) and invisible (*BI*) blinking concerned concentration demand, *BV* imposed markedly less sustained attention than *BI*. Participant feedback clarifies this pattern. P11 observed, "*The highlight feedback indicates I was on target, but I still worried the cursor might be on the edge, seeing it assured me it was fully inside before blinking my eyes.*" This comments suggest that a visible cursor reduces the caution required before issuing a blink confirmation, thereby easing the need for continuous focus.

For dwell, workload, and usability ratings were the lowest among all mechanisms, irrespective of cursor visibility. Nonetheless, comparisons between visible and invisible cursors within the dwell mechanism revealed significant visibility effects across multiple NASA-TLX dimensions—mental demand, temporal demand, perceived performance, effort, frustration, and overall workload, as well as on usability measures of precision, concentration, and frustration. These visibility effects can be attributed to the continuous control nature of dwell, in which users were required to maintain the cursor within the target for a sustained interval. A visible cursor enabled continuous monitoring and correction of cursor position. In contrast, under the invisible cursor condition, unnoticed cursor drift interrupted the selection process, eliciting frustration and confusion as users could not discern when or how drift occurred. Thus, pairing a visible cursor with dwell facilitates cursor maintenance within the target, improving both user experience and performance. These observations mirror our objective results, where the visible cursor outperformed the invisible cursor across all performance metrics (see Sec. 3.6.3).

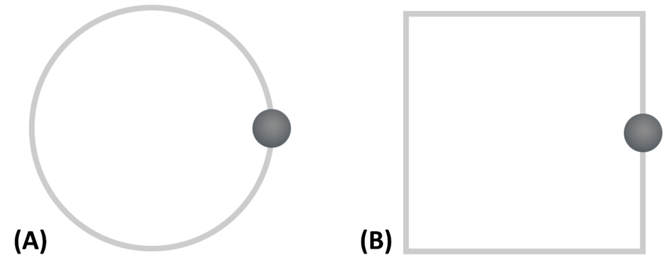


Fig. 5: (A) Circular-shaped path condition. (B) Square-shaped trajectory condition path. At the outset of each trial, participants activated a gray virtual sphere; upon activation, the sphere traversed its predefined trajectory. The movement path remained invisible to participants throughout the experiment.

Taken together, our results answered RQ4: the effect of cursor visibility on user preference depends on the confirmation mechanism. Air tap shows negligible change when the cursor is invisible; blinking experiences a modest reduction in concentration demand with a visible cursor; and dwell selection relies heavily on a visible cursor to reduce workload and enhance usability.

4 User Study 2: Pursuit

The goal of this study is to evaluate and further explore how cursor visibility influences user performance and experience in a pursuit context. We test the following research questions.

- **RQ5**: How does cursor visibility influence users' objective performance in pursuit tasks, and how do these effects vary with target movement speed?
- **RQ6**: Does cursor visibility significantly influence users' subjective experience?

4.1 Participants and Apparatus

A total of 28 participants, aged 18 to 35 years ($M = 22.79$, $SD = 2.85$), were recruited from the local university. Seven of them have participated in our first study. All participants had normal vision, and none of them wore glasses or contact lenses during the experiment. Participants rated their familiarity with VR headsets on a 5-point Likert scale (1 = not familiar; 5 = very familiar), resulting in a mean score of 2.32 ($SD = 1.22$). The same apparatus and experimental development toolkit used in the first study were employed.

4.2 Task

We adapted a pursuit task inspired by prior research [15, 31]. In each trial, participants used their gaze to track a virtual spherical object with a diameter of 2°. The object initially appeared at a fixed location (see Fig. 5) and was activated via an air tap mechanism to ensure fast and consistent trial initiation. As activation served merely to start the tracking sequence and had no bearing on tracking behavior itself, other confirmation mechanisms were not considered. During activation, a boundary highlight provided visual feedback to guide the gaze, and a short beep confirmed that the trial had started. Once activated, the object moved at a constant speed along an invisible path, which followed either a circular or a square trajectory. The trajectory type was randomized for each trial and could be oriented in either a clockwise or counterclockwise direction. Regardless of shape, each trajectory had a perimeter of 30°. At the end of a trial, the object would disappear, and a new object would appear for the next trial. Throughout the study, participants were instructed to maintain their gaze on the moving object as naturally as possible.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2025.3622042

4.3 Design, Procedure, and Measurements

This study adopted a 3×2 within-subjects design with two independent variables: movement speed (V) and cursor visibility (CV). Three movement speeds were tested to represent different difficulty levels ($15°/s$ for easy, $30°/s$ for middle, $45°/s$ for difficult), while cursor visibility used the same visible versus invisible settings as in Study 1. Following Luong et al. [31], target trajectories were rendered as circles or squares to diversify context; however, trajectory shape (S) was not treated as an experimental factor. Each participant experienced the two cursor visibility conditions in alternating order. Within each cursor visibility condition, the combinations of $V\times S$ were presented in a randomized sequence. Each $V\times CV\times S$ combination was repeated five times, resulting in a total of 1,680 trials (28 participants $\times 3V$ $\times 2CV$ $\times 2S$ $\times 5$ repetitions).

We followed the same experimental procedure as in our first study. After completing each cursor visibility condition, participants took a break and filled out questionnaires to report their subjective experience and workload. The entire study lasted approximately 45 minutes per participant.

Objective performance was measured by computing the mean offset and its standard deviation between the pursing object and the cursor's position throughout the trial, sampled at 90 Hz, which represents the *pursuing accuracy* and *stability*, respectively. The same questionnaires used in the first study were employed to evaluate users' workload and utility.

4.4 Results

4.4.1 Results on RQ5: Users' Performance in Pursuit Tasks across Varied Speeds

A two-step pre-processing procedure was conducted to remove outliers. First, trials were discarded if their maximum offset exceeded 15 degrees. Second, trials in which the mean offset or mean deviation exceeded the mean plus or minus three standard deviations were removed. In total, 32 trials (1.9%) were excluded. The remaining data were analyzed using repeated-measures ANOVAs (RM-ANOVAs) to test the main effect of CV and the $V\times CV$ interaction. When normality assumptions were violated, the Aligned Rank Transform (ART) procedure [56] enabled nonparametric testing. Pairwise comparisons were conducted using the Bonferroni–Holm correction with either t-tests or Wilcoxon signed-rank tests, depending on normality of the data.

With respect to pursuing accuracy, the RM-ANOVAs revealed a significant interaction between movement speed and cursor visibility ($F(2,54)=5.773$, $p=.005$, $\eta_p^2=.176$). However, pairwise comparisons showed no significant differences between the two visibility conditions within any given speed level. Instead, significant effects emerged across movement speeds: visible cursors yielded greater offsets at the fastest speed compared to both low ($Z=4.59$, $\Delta=0.78$, $p<0.001$, $r=0.867$) and mid-level speeds ($Z=4.34$, $\Delta=0.46$, $p<0.001$, $r=0.820$), and mid-level speeds also produced larger offsets than low speeds ($Z=3.38$, $\Delta=0.32$, $p=0.008$, $r=0.639$).

Regarding stability, the RM-ANOVAs revealed neither a main effect of cursor visibility nor any interaction with it.

4.4.2 Results on RQ6: Subjective Experience under Different Cursor Visibility in Pursuit Tasks

For subjective measures, the same analysis procedures as those used for the subjective results in Study 1 were applied. Among the measures, only naturalness exhibited a significant difference in usability, with the invisible cursor yielding a significantly higher score than the visible cursor ($t(27)=3.14$, $\Delta=0.67$, $p=0.023$, $d_z=0.521$).

4.5 Discussion

4.5.1 Discussion of RQ5: Performance Remains Comparable Across Visibility Conditions

Our results indicate no significant main effect in pursuing accuracy or stability between the two cursor visibility conditions. Nevertheless, further analysis of pursuing accuracy revealed a significant interaction between movement speed and cursor visibility. Based on the pairwise comparisons, these results can be interpreted as movement speed exerting a stronger influence on pursuing accuracy under the visible cursor condition, where offsets increased as speed rose and reached their maximum at the fastest level. In contrast, this trend was less pronounced under the invisible cursor condition. One possible explanation for this pattern is that a visible cursor provides an additional spatial reference at lower speeds, supporting more precise alignment. At higher speeds, however, the same feedback becomes distracting, leading to disrupted tracking and larger offsets. Participants' interview responses echoed this trend: at low speeds, the visible cursor was considered helpful for alignment, whereas at higher speeds it was often described as redundant or even distracting, competing with the moving target for attention. Conversely, under the invisible condition, the absence of continuous cursor feedback encouraged participants to rely more consistently on the target itself, which may explain why speed exerted a weaker influence on accuracy.

4.5.2 Discussion of RQ6: Visibility Affects Perceived Naturalness

In response to RQ6, significant effects were observed only for the naturalness usability dimension; no other subjective measures in the pursuit task showed significant effects of cursor visibility. Specifically, most participants favored the invisible cursor, citing greater naturalness and fewer visual distractions. This aligns with our objective explanation that the visible cursor may become distracting under high-demand pursuit conditions. As P18 put it, "*The invisible cursor doesn't pull my attention away, whereas the visible one feels distracting and unnatural as it circles the target.*"

This preference reflects the pursuit task's main demand, keeping gaze motion synchronized with the target, rather than pinpoint positional accuracy [15, 39]. In that case, the visible cursor's precision benefit is outweighed by its visual clutter and distraction. Consequently, the change in task reshaped participants' visibility preferences: P7 and P25, who preferred a visible cursor in Study 1's selection task, now favored the invisible cursor, noting, "*Precise positioning mattered in Study 1, but not here, so the invisible cursor feels better.*"

5 Design Implications

Based on the results of the two studies, we distilled the following Design Implications (Noted as **DI#**).

DI1: Air Tap Delivers Consistently Superior Performance. Although dwell achieved the highest throughput and success rate under the visible cursor condition, air tap consistently received superior subjective ratings across all measures; it even outperformed dwell in throughput under the invisible cursor condition. These findings highlight air tap as a reliable and user-preferred confirmation mechanism that balances speed, accuracy, and experience across varying visibility settings in VR gaze-based selection tasks. Designers can thus rely on air tap to deliver robust performance and enhanced user experience within different selection tasks, ensuring efficient and reliable selections in VR.

DI2: Use Dwell Confirmation When Hands-Free and High Reliability Are Required. Dwell-based selection, which requires users to maintain their gaze on a target for a fixed duration, consistently yields the highest success rate across cursor visibility conditions compared to other mechanisms. However, it frequently leads to prolonged selection times and increased user frustration, especially under high-difficulty conditions such as small target sizes, where it becomes difficult to maintain the gaze point within the target area long enough to trigger confirmation reliably. Therefore, designers should use dwell-based confirmation cautiously and reserve it for situations requiring hands-free interaction and where a high success rate is needed for user operations.

DI3: Air Tap Performance Remains Stable Regardless of Cursor Visibility. Our findings reveal that, with consistent feedback such as boundary highlighting, the effectiveness of the air tap remains unchanged whether the cursor is visible or not. In this case, designers can choose a visible or an invisible cursor based on aesthetic or interface clarity considerations without compromising the performance of air tap interactions.

DI4: Maintain Visible Cursor When Using Blinking for Confirmation. Employing a blinking confirmation mechanism with a continuously visible cursor markedly enhances selection accuracy across almost all task conditions, despite increased selection times under low-challenge conditions (e.g., large width or small amplitude) and higher reentry rates under high-challenge conditions (e.g., small width or large amplitude). These performance costs remain far outweighed by the substantial gains in precision, andgiven users' preference for visible cursor-we therefore recommend maintaining cursor visibility whenever blinking serves as the confirmation mechanism.

DI5: Use a Visible Cursor to Improve Dwell-Based Selections. When dwell is used for confirmation, incorporating a visible cursor provides valuable real-time cues, enabling users to verify their gaze alignment with the target. This improves selection accuracy and reduces cognitive workload. We therefore recommend integrating a visible cursor into the dwell mechanism to enhance both performance and user experience.

DI6: Invisible Cursor Improves Comfort in Gaze-Based Pursuit Task. For pursuit tasks involving continuous gaze tracking of moving targets, our findings suggest that an invisible cursor provides a more natural and less visually distracting experience. We therefore recommend using an invisible cursor when the task emphasizes motion correlation over precise target acquisition.

DI7: Provide Supplementary Feedback to Support Gaze-Based Interaction. Our findings diverge from prior research on selection tasks [2], likely due to the inclusion of supplementary feedback mechanisms, such as boundary highlighting, which were not present in earlier studies. These results underscore the importance of providing strong, task-relevant feedback to help users maintain awareness of their gaze position and selection status, regardless of cursor visibility. We therefore recommend integrating explicit visual feedback into gaze-based interfaces to improve performance and user confidence.

DI8: Adapt Cursor Visibility to Match Task Demands. Our findings indicate that the effectiveness of cursor visibility depends on the nature of the interaction task. For discrete selection tasks, a visible cursor improves precision, particularly when using confirmation mechanisms such as blinking or dwell. In contrast, for pursuit tasks that require continuous gaze tracking, an invisible cursor offers a more natural experience by minimizing visual distractions. Based on these insights, we recommend adapting cursor visibility to task requirements: use a visible cursor with supplementary feedback for accuracy-focused tasks, and an invisible cursor for tracking-oriented tasks where seamless, distraction-free interaction is preferred.

6 Limitations and Future Work

This section outlines the limitations of this work and suggests directions for future research. Recognizing these boundaries is important for interpreting our findings and for informing subsequent investigations. First, while our study thoroughly evaluated the effects of the cursor visibility in VR systems, it focused solely on two fundamental and typical tasks: selection and pursuit. Our findings suggest that different tasks have unique requirements, which, in turn, affect how the cursor visibility performs. Therefore, future studies are recommended to investigate more complex application contexts to gain a deeper understanding of the effects of cursor visibility. Second, our evaluation relied on the most popular VR headset equipped with a built-in eye-tracking device. Although this device is representative in terms of eye-tracking accuracy [4], further research can explore how variations across different devices and their respective accuracies might influence user performance and overall experience with various cursor metaphors.

7 Conclusion and Summary of Findings

This paper presents two user studies examining the impact of cursor visibility (visible vs. invisible) on performance and user experience across gaze-based selection and pursuit tasks in virtual reality (VR) scenarios.

In the pointing-selection study, cursor visibility fundamentally reshaped the relative performance of three confirmation mechanisms. With a visible cursor, dwell achieved the highest success rate and throughput, and air tap clearly outperformed blinking. When the cursor was removed, this ranking reversed: air tap emerged as the top performer, while both dwell and blinking declined in performance. These visibility effects were further modulated by movement amplitude and target width, with each mechanism showing a distinct response pattern: air tap was unaffected by visibility, blinking showed a task challenge-scaled effect from visibility, and dwell consistently improved with cursor presence. Subjective ratings mirrored these trends—neutral for air tap, mildly positive for blinking, and strongly positive for dwell.

By contrast, the pursuit task study revealed a different pattern. A visible cursor showed no overall performance advantage; offsets rose more substantially with speed under the visible condition, whereas the invisible cursor was less affected and was also rated as more natural.

Together, these findings emphasize the importance of adapting cursor visibility to the specific demands of the task. Based on our empirical results, we propose eight design implications that illustrate how cursor visibility and confirmation mechanisms interact. These implications highlight not only the important role of the cursor visibility but also the necessity of gaining deeper insights into its effects across diverse use cases.

Acknowledgments

References

[1] ISO/TS 9241-411:2012 Ergonomics of human-system interaction — Part 411: Evaluation methods for the design of physical input devices, 2012. Accessed: 2025-04-16. 3

[2] T. S. M. Ajoy S. Fernandes, Immo Schütz and M. J. Proulx. Gaze inputs for targeting: The eyes have it, not with a cursor. *International Journal of Human–Computer Interaction*, 0(0):1–19, 2025. doi: 10.1080/10447318.2025.2453966 1, 2, 3, 8, 11

[3] R. Alonso, M. Causse, F. Vachon, R. Parise, F. Dehais, and P. Terrier. Evaluation of head-free eye tracking as an input device for air traffic control. *Ergonomics*, 56(2):246–255, 2013. doi: 10.1080/00140139.2012.744473 2

[4] S. Aziz, D. J. Lohr, L. Friedman, and O. Komogortsev. Evaluation of eye tracking signal quality for virtual reality applications: A case study in the meta quest pro. In *Proceedings of the 2024 Symposium on Eye Tracking Research and Applications*, ETRA '24, article no. 7, 8 pages. Association for Computing Machinery, New York, NY, USA, 2024. doi: 10.1145/3649902.3653347 11

[5] G. R. Barnes. *Ocular pursuit movements*, pp. 116–132. Oxford University Press, 2012. doi: 10.1093/oxfordhb/9780199539789.013.0007 2

[6] G. R. Barnes. Rapid learning of pursuit target motion trajectories revealed by responses to randomized transient sinusoids. *Journal of Eye Movement Research*, 5(3), Jun. 2012. doi: 10.16910/jemr.5.3.4 2

[7] M. Bernhard, E. Stavrakis, M. Hecher, and M. Wimmer. Gaze-to-object mapping during visual search in 3d virtual environments. *ACM Transactions on Applied Perception*, 11(3):17, aug 2014. doi: 10.1145/2644812 2

[8] E. Birmingham, W. F. Bischof, and A. Kingstone. Gaze selection in complex social scenes. *Visual Cognition*, 16(2-3):341–355, 2008. doi: 10.1080/13506280701434532 1, 2

[9] J. Blattgerste, P. Renner, and T. Pfeiffer. 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*, COGAIN '18, article no. 1, 9 pages. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3206343.3206349 1, 2

[10] E. Cáceres, M. Carrasco, and S. Ríos. Evaluation of an eye-pointer interaction device for human-computer interaction. *Heliyon*, 4:e00574, 2018. doi: 10.1016/j.heliyon.2018.e00574 1

[11] B. David-John, C. Peacock, T. Zhang, T. S. Murdison, H. Benko, and T. R. Jonker. Towards gaze-based prediction of the intent to interact in virtual reality. *ACM Symposium on Eye Tracking Research and Applications*, 2021. doi: 10.1145/3448018.3458008 2

[12] T. J. Dube, Y. Ren, H. Limerick, I. S. MacKenzie, and A. S. Arif. Push, tap, dwell, and pinch: Evaluation of four mid-air selection methods augmented with ultrasonic haptic feedback. *Proc. ACM Hum.-Comput. Interact.*, 6(ISS), article no. 565, 19 pages, Nov. 2022. doi: 10.1145/3567718 2, 8

[13] A. T. Duchowski. Gaze-based interaction: A 30 year retrospective. *Computers & Graphics*, 73:59–69, 2018. doi: 10.1016/j.cag.2018.04.002 2

[14] S. Esaki, Y. Ebisawa, A. Sugioka, and M. Konishi. Quick menu selection using eye blink for eye-slaved nonverbal communicator with video-based eye-gaze detection. In *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 'Magnificent Milestones and Emerging Opportunities in Medical Engineering' (Cat. No.97CH36136)*, vol. 5, pp. 2322–2325 vol.5, 1997. doi: 10.1109/IEMBS.1997.758829 2

[15] A. Esteves, E. Velloso, A. Bulling, and H. Gellersen. Orbits: Gaze interaction for smart watches using smooth pursuit eye movements. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, UIST '15, 10 pages, p. 457–466. Association for Computing Machinery, New York, NY, USA, 2015. doi: 10.1145/2807442.2807499 2, 9, 10

[16] A. M. Feit, S. Williams, A. Toledo, A. Paradiso, H. Kulkarni, S. Kane, and M. R. Morris. 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*, CHI '17, 13 pages, p. 1118–1130. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3025453.3025599 1, 2, 4

[17] A. S. Fernandes, T. S. Murdison, and M. Proulx. Leveling the playing field: A comparative reevaluation of unmodified eye tracking as an input and interaction modality for vr. *IEEE Transactions on Visualization and Computer Graphics*, 29:2269–2279, 2023. doi: 10.1109/TVCG.2023.3247058 1, 2

[18] D. Fono and R. Vertegaal. Eyewindows: evaluation of eye-controlled zooming windows for focus selection. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, 10 pages, p. 151–160. Association for Computing Machinery, New York, NY, USA, 2005. doi: 10.1145/1054972.1054994 2

[19] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, pp. 904–908. Sage publications Sage CA: Los Angeles, CA, 2006. 4

[20] F. Heinrich, F. Huettl, G. Schmidt, M. Paschold, W. Kneist, T. Huber, and C. Hansen. Holopointer: a virtual augmented reality pointer for laparoscopic surgery training. *International Journal of Computer Assisted Radiology and Surgery*, 16(1):161–168, Jan 2021. doi: 10.1007/s11548-020-02272-2 1, 2

[21] Herlina, S. Wibirama, and I. Ardiyanto. Similarity measures of object selection in interactive applications based on smooth pursuit eye movements. *2018 International Conference on Information and Communications Technology (ICOIACT)*, pp. 639–644, 2018. doi: 10.1109/ICOIACT.2018.8350701 2

[22] X. Hu, Y. Zhang, Y. Wei, L. Zhang, Y. Li, W. Stuerzlinger, and H.-N. Liang. Exploring and modeling the effects of eye-tracking accuracy and precision on gaze-based steering in virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–11, 2025. doi: 10.1109/TVCG.2025.3616824 1, 2

[23] T. Isomoto, S. Yamanaka, and B. Shizuki. Dwell selection with ml-based intent prediction using only gaze data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 6(3), article no. 120, 21 pages, Sept. 2022. doi: 10.1145/3550301 2

[24] X. Jin, S. Chai, J. Tang, X. Zhou, and K. Wang. Eye-tracking in ar/vr: A technological review and future directions. *IEEE Open Journal on Immersive Displays*, 1:146–154, 2024. doi: 10.1109/OJID.2024.3450657 2

[25] W. jun Hou and X. lin Chen. Comparison of eye-based and controller-based selection in virtual reality. *International Journal of Human–Computer Interaction*, 37:484 – 495, 2020. doi: 10.1080/10447318.2020.1826190 2

[26] M. Land and B. Tatler. *Looking and Acting: Vision and Eye Movements in Natural Behaviour*. Oxford University Press, 2009. doi: 10.1093/acprof:oso/9780198570943.001.0001 3

[27] H. S. Lee, F. Weidner, L. Sidenmark, and H. Gellersen. Snap, pursuit and gain: Virtual reality viewport control by gaze. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, article no. 375, 14 pages. Association for Computing Machinery, New York, NY, USA, 2024. doi: 10.1145/3613904.3642838 2

[28] T. Lei, J. Chen, J. Chen, and B. Liu. Modeling the gaze point distribution to assist eye-based target selection in head-mounted displays. *Neural Computing and Applications*, 35(36):25069–25081, 2023. doi: 10.1007/s00521-023-08705-8 2

[29] X. Lu, D. Yu, H.-N. Liang, and J. Goncalves. itext: Hands-free text entry on an imaginary keyboard for augmented reality systems. In *The 34th Annual ACM Symposium on User Interface Software and Technology*, UIST '21, 11 pages, p. 815–825. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3472749.3474788 1, 2, 3

[30] X. Lu, D. Yu, H.-N. Liang, W. Xu, Y. Chen, X. Li, and K. Hasan. Exploration of hands-free text entry techniques for virtual reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 344–349, 2020. doi: 10.1109/ISMAR50242.2020.00061 2

[31] T. Luong, Y. F. Cheng, M. Mobus, 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 & Computer Graphics*, 29(11):4633–4643, Nov. 2023. doi: 10.1109/TVCG.2023.3320211 9, 10

[32] O. H.-M. Lutz, A. C. Venjakob, and S. Ruff. Smoovs: Towards calibration-free text entry by gaze using smooth pursuit movements. *Journal of Eye Movement Research*, 8(1):2, 2015. doi: 10.16910/jemr.8.1.2 2

[33] M. N. Lystbæk, K. Pfeuffer, J. E. S. Grønbæk, and H. Gellersen. Exploring gaze for assisting freehand selection-based text entry in ar. *Proc. ACM Hum.-Comput. Interact.*, 6(ETRA), article no. 141, 16 pages, May 2022. doi: 10.1145/3530882 2

[34] M. N. Lystbæk, K. Pfeuffer, J. E. S. Grønbæk, and H. Gellersen. Exploring gaze for assisting freehand selection-based text entry in ar. *Proc. ACM Hum.-Comput. Interact.*, 6(ETRA), article no. 141, 16 pages, May 2022. doi: 10.1145/3530882 2

[35] C. Mei, B. T. Zahed, L. Mason, and J. Ouarles. Towards joint attention training for children with asd - a vr game approach and eye gaze exploration. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 289–296, 2018. doi: 10.1109/VR.2018.8446242 2

[36] Microsoft. Gaze and dwell for eyes, 2024. Accessed: 2024-12-15. 2

[37] T. Moscovich and J. F. Hughes. Multi-finger cursor techniques. In *Proceedings of Graphics Interface 2006*, GI '06, 7 pages, p. 1–7. Canadian Information Processing Society, CAN, 2006. 1, 2

[38] A. K. Mutasim, A. U. Batmaz, and W. Stuerzlinger. 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*, ETRA '21 Short Papers, article no. 15, 7 pages. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3448018.3457998 2, 8

[39] O. Namnakani, Y. Abdrabou, J. Grizou, A. Esteves, and M. Khamis. Comparing dwell time, pursuits and gaze gestures for gaze interaction on handheld mobile devices. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, article no. 258, 17 pages. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544548.3580871 2, 8, 10

[40] Y. T. Paulus and G. B. Remijn. Usability of various dwell times for eye-gaze-based object selection with eye tracking. *Displays*, 67:101997, 2021. doi: 10.1016/j.displa.2021.101997 2

[41] K. Pfeuffer, M. Vidal, J. Turner, A. Bulling, and H. Gellersen. Pursuit calibration: making gaze calibration less tedious and more flexible. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST '13, 10 pages, p. 261–270. Association for Computing Machinery, New York, NY, USA, 2013. doi: 10.1145/2501988.2501998 2

[42] T. Piumsomboon, G. Lee, R. W. Lindeman, and M. Billinghurst. Exploring natural eye-gaze-based interaction for immersive virtual reality. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 36–39, 2017. doi: 10.1109/3DUI.2017.7893315 2, 4

[43] Y. Qian and R. J. Teather. The eyes don't have it: an empirical comparison of head-based and eye-based selection in virtual reality. *Proceedings of the 5th Symposium on Spatial User Interaction*, 2017. doi: 10.1145/3131277.3132182 1, 2

[44] J. Ross, M. Morrone, M. Goldberg, and D. Burr. Changes in visual perception at the time of saccades. *Trends in Neurosciences*, 24:113–121, 2001. doi: 10.1016/S0166-2236(00)01685-4 2

[45] R. V. S, A. S, A. R, N. V, and G. Prabhakar. To show or hide - investigating the effect of cursor visibility in gaze-controlled interfaces. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 2024. doi: 10.1145/3613905.3651005 1, 2, 8

[46] R. Shi, Y. Wei, X. Qin, P. Hui, and H.-N. Liang. Exploring gaze-assisted and hand-based region selection in augmented reality. *Proc. ACM Hum.-Comput. Interact.*, 7(ETRA), article no. 160, 19 pages, May 2023. doi: 10.1145/3591129 1

[47] L. E. Sibert and R. J. K. Jacob. Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '00, 8 pages, p. 281–288. Association for Computing Machinery, New York, NY, USA, 2000. doi: 10.1145/332040.332445 1, 2

[48] L. Sidenmark, M. Parent, C.-H. Wu, J. Chan, M. Glueck, D. J. Wigdor, T. Grossman, and M. Giordano. Weighted pointer: Error-aware gaze-based interaction through fallback modalities. *IEEE Transactions on Visualization and Computer Graphics*, 28:3585–3595, 2022. doi: 10.1109/TVCG.2022.3203096 1

[49] L. Sidenmark, F. Prummer, J. Newn, and H. Gellersen. Comparing gaze, head and controller selection of dynamically revealed targets in head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 29(11):4740–4750, 2023. doi: 10.1109/TVCG.2023.3320235 2

[50] J. Software. Het selecteren via knipperen bij oogsturing instellen, 2024. Accessed: 2024-12-15. 2

[51] N. A. Suwadi, M. C. Lam, and N. A. A. Majid. Evaluation on different object selection visual feedback in collaborative augmented reality. *International Journal of Human–Computer Interaction*, 41:7135 – 7151, 2024. doi: 10.1080/10447318.2024.2389727 3

[52] V. Tanriverdi and R. Jacob. Interacting with eye movements in virtual environments. *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 2000. doi: 10.1145/332040.332443 2

[53] A. S. Taylor. After interaction. *Interactions*, 22:48 – 53, 2015. doi: 10.1145/2809888 1

[54] M. Vidal, A. Bulling, and H. Gellersen. Pursuits: Spontaneous interaction with displays based on smooth pursuit eye movement and moving targets. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (Zurich, Switzerland)*, UbiComp '13, pp. 439–448. Association for Computing Machinery, New York, NY, USA, 2013. doi: 10.1145/2493432.2493477 2

[55] Y. Wei, R. Shi, D. Yu, Y. Wang, Y. Li, L. Yu, and H.-N. Liang. Predicting gaze-based target selection in augmented reality headsets based on eye and head endpoint distributions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, article no. 283, 14 pages. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544548.3581042 1, 2, 3, 4, 8

[56] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, 4 pages, p. 143–146. Association for Computing Machinery, New York, NY, USA, 2011. doi: 10.1145/1978942.1978963 5, 10

[57] D. Yu, R. Desai, T. Zhang, H. Benko, T. R. Jonker, and A. Gupta. Optimizing the timing of intelligent suggestion in virtual reality. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, UIST '22, article no. 6, 20 pages. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3526113.3545632 3

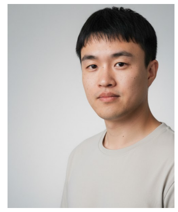
[58] D. Yu, H.-N. Liang, X. Lu, K. Fan, and B. Ens. Modeling endpoint distribution of pointing selection tasks in virtual reality environments. *ACM Trans. Graph.*, 38(6), article no. 218, 13 pages, Nov. 2019. doi: 10.1145/3355089.3356544 1

[59] D. Yu, X. Lu, R. Shi, H.-N. Liang, T. Dingler, E. Velloso, and J. Goncalves. Gaze-supported 3d object manipulation in virtual reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, article no. 734, 13 pages. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3411764.3445343 1, 2

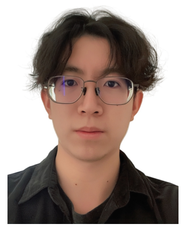
[60] A. Zenner, K. Ullmann, O. Ariza, F. Steinicke, and A. Krüger. Induce a blink of the eye: Evaluating techniques for triggering eye blinks in virtual reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, article no. 256, 12 pages. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544548.3580888 2

[61] A. Zhu, S. Cheng, and J. Fan. Eye tracking and gesture based interaction for target selection on large displays. *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, 2018. doi: 10.1145/3267305.3267607 2

Author Biographies



Yushi Wei is a Ph.D. candidate at the Hong Kong University of Science and Technology (Guangzhou), China. His research interests include human-computer interaction, virtual and augmented reality, and interaction design.



Rongkai Shi received the B.S. and Ph.D. degrees from the University of Liverpool, Liverpool, U.K., and Xi'an Jiaotong-Liverpool University, China. He is currently a postdoctoral researcher at the Hong Kong University of Science and Technology (Guangzhou), China. His research interests include human-computer interaction, extended reality, and interaction design.

Sen Zhang is currently an undergraduate student in the School of Advanced Technology at Xi'an Jiaotong-Liverpool University (XJTLU), China. His academic interests include virtual reality, human–computer interaction, and computer graphics.

Anil Ufuk Batmaz received his PhD in Biomedical Engineering from Strasbourg University, France. He is currently an Assistant Professor in the Department of Computer Science and Software Engineering at Concordia University, Gina Cody School of Engineering and Computer Science, in Montreal, Canada. He leads the "EXtended Reality and Interaction Technologies" (EXIT) Lab. His research interests include virtual reality, augmented reality, training systems and simulators, human-computer interaction, and human–AI interaction.

Pan Hui is Chair Professor of Computational Media and Arts and Director of the Center for Metaverse and Computational Creativity at HKUST (Guangzhou). A globally recognized expert in ubiquitous computing and AR/VR, he is a Fellow of the Royal Academy of Engineering and IEEE, and a Member of Academia Europaea. He has held prominent roles in both academia and industry across Europe and Asia, including Distinguished Scientist at Deutsche Telekom Laboratories and Nokia Chair in Data Science at the University of Helsinki. He received his Computer Science PhD from the University of Cambridge.

Hai-Ning Liang is an Associate Professor with the Computational Media and Arts Thrust at the Hong Kong University of Science and Technology (Guangzhou), China. He received his PhD in computer science from the University of Western Ontario, Canada. His research focuses on the design and evaluation of novel interactions and applications for virtual and mixed reality, gaming, and visualization technologies.