FlashCue: Momentary Displayed Visual Notifications in Extended Reality

Atsushi Orii
IPLAB
University of Tsukuba
Tsukuba, Ibaraki, Japan
orii@iplab.cs.tsukuba.ac.jp

Yusuke Ashizawa
IPLAB
University of Tsukuba
Tsukuba, Ibaraki, Japan
ashizawa@iplab.cs.tsukuba.ac.jp

Myungguen Choi
IPLAB
University of Tsukuba
Tsukuba, Ibaraki, Japan
choi@iplab.cs.tsukuba.ac.jp

Buntarou Shizuki
IPLAB
University of Tsukuba
Tsukuba, Ibaraki, Japan
shizuki@cs.tsukuba.ac.jp

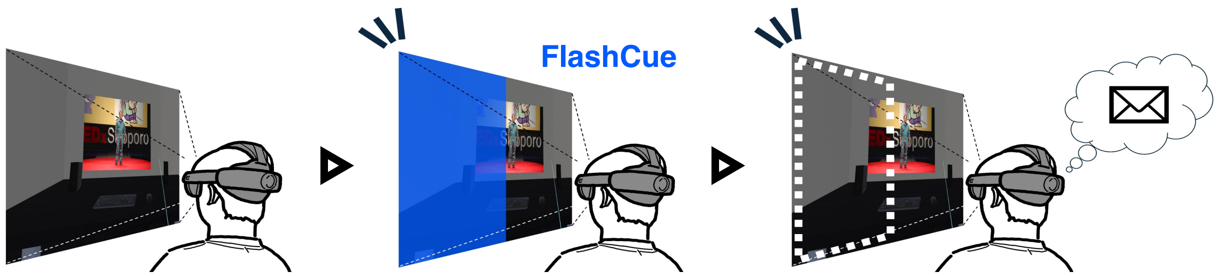

Figure 1: FlashCue is a notification for extended reality (xR) devices that momentarily displays visual information, thereby enabling notifications to be delivered to the user without obstructing their field of view. This information is delivered through the color and position of the notification, enabling users to quickly and efficiently read simple visual information.

Abstract

We present FlashCue: a visual notification method designed to deliver simple cues in extended reality environments. FlashCue provides the user with visual stimuli for an extremely short duration, allowing notifications to be delivered without obstructing the user's field of view. By varying its color and display position, FlashCue can present different types of notifications. In this study, we examined the range of notifications that FlashCue can deliver. Our results indicate that, with its display duration of 1/90 second, FlashCue can deliver 32 distinct notifications by combining four colors with eight display positions.

CCS Concepts

• **Human-centered computing** → **Virtual reality**.

Keywords

Notifications, Virtual Reality, Augmented Reality, Head Mounted Display

OZCHI '25, Sydney, Australia

ACM ISBN 979-8-4007-2016-1/25/11
https://doi.org/10.1145/3764687.3769908

ACM Reference Format:
Atsushi Orii, Yusuke Ashizawa, Myungguen Choi, and Buntarou Shizuki. 2025. FlashCue: Momentary Displayed Visual Notifications in Extended Reality. In *37th Australian Conference on Human-Computer Interaction (HCI) (OZCHI '25), November 29–December 03, 2025, Sydney, Australia.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3764687.3769908

1 Introduction

A notification is a visual, auditory, or haptic alert designed to capture users' attention and proactively deliver information [5, 19, 36]. Users encounter numerous notifications daily, such as those from smartphone applications and email clients [33, 35]. These notifications are traditionally delivered on devices such as smartphones [17, 29] and desktop computers [20, 33, 47]. However, given the advancement and widespread adoption of smart glasses and head-mounted displays (HMDs), new notification methods have been proposed for extended reality (xR) environments, which include augmented reality (AR) and virtual reality (VR).

While visual notifications are commonly used in xR environments, they can be unsuitable in certain situations because they block the user's field of view (FOV). For example, visual notifications delivered while a user is walking or driving can potentially lead to collisions. Non-visual notifications that rely on sound and haptic feedback are often employed to address this limitation. However, audio notifications can be missed by the user in noisy environments and are often not socially acceptable in public spaces. Furthermore, vibration-based notifications are limited in variety and unable to deliver large amounts of information [43, 46].

In this article, we propose *FlashCue*, a visual notification method that presents simple visual cues in the user's FOV for an extremely brief duration. Because FlashCue appears momentarily, it minimizes disruptions to the user's FOV. Thus, even though FlashCue relies on visual information, it can function similarly to non-visual notification methods. Furthermore, by varying its color and layout, a wide range of notifications can be delivered using FlashCue and no other notification systems. For example, briefly displaying a blue FlashCue on the left side of the screen informs the user that they have received an email from their family (Fig. 1). In this way, adjusting FlashCue's position or color makes it possible to encode information such as the sender or content type of an email.

To the best of our knowledge, no prior studies have examined notification methods that use momentary visual stimuli; therefore, little is known about the optimal display duration or placement for such stimuli. Therefore, we sought to determine the appropriate display time for FlashCue and the number of notifications it can convey. In Study #1, we examined the recognizable display duration and opacity of FlashCue. Results showed that participants could recognize FlashCue even with a very brief display duration (1/90 s $\approx$ 0.011 s) and preferred an opacity of 75%. Study #2 assessed whether participants could recognize FlashCues while performing other tasks. The results indicated that participants were able to distinguish 32 types of FlashCue with an error rate of 5.73% while solving a calculation problem or watching videos.

2 Related Work

In this section, we discuss notification methods in xR environments and the human ability to recognize rapidly presented visual information.

2.1 Notification Method in xR Environments

Since xR technologies enable the seamless integration of digital content into users' physical surroundings, researchers have investigated notifications designed specifically for xR conditions, such as walking [4, 12, 45], cycling [28], remote collaboration [3], or face-to-face conversations [15, 22, 41].

The presentation of visual notifications in xR environments can obstruct the user's FOV, causing visual occlusion and distraction, resulting in interference with user behavior [6, 28, 37, 41]. Research has shown that the presentation and placement of xR notifications significantly influence response time, noticeability, distraction, and intrusiveness [8, 42]. Various methods have been proposed to address these challenges, including adjusting the display position of notifications [31, 39, 45], metaphors of notifications [38], and presenting notifications at opportune moments [2, 27]. For instance, Shin et al. [45] investigated the opacity and position of notifications in video see-through AR in terms of their noticeability and users' preferences while walking.

2.2 Rapid Visual Recognition

The human recognition process can occur without focused attention and often involves unconscious information processing. This ability enables humans to rapidly classify images presented for extremely short durations. Thorpe et al. [48] were the first to show that humans can accurately classify images presented to them for periods as short as 20 ms. Subsequent research confirmed that even a single monitor refresh (10–40 ms) can provide humans with sufficient information for advanced scene analysis [13, 14, 25, 26]. Similarly, Potter et al. [40] found that humans could recognize images presented for approximately 13 ms, even when such images were shown as part of a continuous sequence of 12 images. In addition, Lanfranco et al. [30] investigated the minimum presentation duration required for image recognition, finding that the minimum time needed to achieve performance significantly above chance was 1.7 ms.

While previous studies have established the fundamental temporal limits of human visual recognition, xR-specific factors can further constrain performance. Gabbard et al. [9–11] reported that lighting conditions, background complexity, and frequent focal changes can degrade visual performance and increase attentional costs in xR environments. Building on these findings, our study examines the recognition of very briefly presented visual information in xR environments.

3 Study #1

The extent to which people can recognize brief visual information presented for an extremely short duration in xR environments remains unclear. In this study, we examined (1) whether users can recognize instantaneously presented visual information and (2) the appropriate display duration and opacity of FlashCue. The study tasks were performed in VR environments.

3.1 Participants and Apparatus

Ten participants (all male; mean age = 23.5 years, SD = 1.0 years; IDs: P1–P10) from the authors' laboratory were recruited for the study. All participants had normal or corrected-to-normal vision.

We used the Meta Quest 3, along with two of its controllers, as the HMD in this study. According to the HMD's specifications, the visible FOV is 110° horizontally and 96° vertically. The application used in this study was developed using Unity (version 2022.3.4f1). Although the HMD's refresh rate was 120 Hz, the application ran at 90 Hz because Unity's update function operated at this refresh rate.

3.2 Design

We employed a within-participant design with two independent variables (Fig. 2):

- *Duration*: 1/90 s ($\approx$ 0.011 s), 2/90 s ($\approx$ 0.022 s), 3/90 s ($\approx$ 0.033 s)
- *Opacity*: 100%, 75%, 50%, 25%

Duration refers to the display time of FlashCue. The maximum frame rate of the Unity application is 90 Hz—the limit of Unity's update function. Although the refresh rate of the HMD exceeds 90 Hz, the application itself does not operate above 90 Hz. Therefore, we selected three values corresponding to the shortest frame intervals that could be presented using Unity's update function. *Opacity* refers to the transparency level of FlashCue (Figure 2b). Fully opaque cues completely obstruct the user's vision and may not be optimal in this study. Therefore, *Opacity* was included as a parameter to explore users' preferences regarding the transparency of FlashCue.

The participants completed three sessions with 48 trials (4 FlashCue colors × 3 *Durations* × 4 *Opacities*). Four colors—red, blue, yellow, and sky blue—were selected from a colorblind-safe palette [21]. The order of the 48 trials was randomized. The participants also completed the three sessions in each of two rooms. Both rooms were selected based on the study by Li et al. [32]. The office room featured cool lighting, while the living room featured warm lighting. The order of room presentation was counterbalanced. Two room types and four FlashCue colors were included in the design to minimize potential biases in recognition accuracy due to background characteristics or color differences. In total, we collected 288 data points per participant (2 rooms × 3 sessions × 48 trials), which amounted to 2,880 data points across all ten participants.

3.3 Procedure and Task

First, the participants were seated and asked to complete a questionnaire collecting personal information, such as their name and gender. After completing the questionnaire, they were given detailed instructions about the task. Subsequently, the participants wore the HMD and held the controller.

The participants were tasked to press the controller's trigger when FlashCue was recognized. Once the task began, one of the rooms was displayed. After the participants pressed the trigger to indicate the start of the task, a FlashCue was displayed within 2–5 seconds. The FlashCues of this study covered the participants' entire FOV. The participants were instructed to press the controller's trigger as soon as they saw a FlashCue. After each FlashCue was displayed, the next FlashCue appeared within 2–5 seconds. The frequency of presentation used in this study was considered to have an extremely low risk of inducing epilepsy [24]. This procedure was repeated until FlashCue was displayed 48 times, which constituted one session. A break of at least one minute was provided after each session.

After completing three consecutive sessions in one room, the participants completed another three sessions in the other room. After completing all trials, the participants were interviewed and asked to indicate their preferences for *Duration* and *Opacity*. Each study session took approximately 60 minutes.

3.4 Metrics

The main dependent variable was the recognition error rate, defined as the proportion of trials in which FlashCue was not recognized. This was calculated by dividing the number of unrecognized trials by the total number of trials. A trial was deemed unrecognized if the trigger was not pressed within two seconds after a FlashCue was displayed, which was based on the minimum time interval before the next FlashCue was presented. We also collected the participants' preferences for each *Duration* and *Opacity*.

3.5 Results

3.5.1 Recognition Error Rate. To analyze the recognition error rate, a nonparametric aligned rank transformation (ART) method [16, 44, 49] was employed followed by a two-way repeated measures ANOVA with *Duration* and *Opacity* as factors. Post-hoc analyses for within-factor comparisons were performed using ART-C [7] followed by Holm correction [18].

The recognition error rates are provided in Fig. 3. Specifically, across all trials, the recognition error rate of FlashCue was 3.92%. We observed that *Duration* ($F_{2,2859} = 59.1, p < .01, \eta_p^2 = .039$) and *Opacity* ($F_{3,2859} = 99.74, p < .01, \eta_p^2 = .095$) had significant effects on the recognition error rate. Significant interactions were also observed for *Duration* × *Opacity* ($F_{6,2859} = 42.4, p < .01, \eta_p^2 = .082$).

3.5.2 Participants' Preference. The average preference ranks for each *Duration* (1/90 s, 2/90 s, and 3/90 s) were 2.2 (1st: 4, 2nd: 0, 3rd: 6), 1.6 (1st: 4, 2nd: 6, 3rd: 0), and 2.2 (1st: 2, 2nd: 4, 3rd: 4), respectively.

The average preference ranks for each *Opacity* (100%, 75%, 50%, 25%) were 2.5 (1st: 2, 2nd: 4, 3rd: 1, 4th: 3), 1.9 (1st: 4, 2nd: 3, 3rd: 3, 4th: 0), 2.4 (1st: 2, 2nd: 2, 3rd: 6, 4th: 0), and 3.2 (1st: 2, 2nd: 1, 3rd: 0, 4th: 7), respectively.

3.5.3 Participants' Feedback. Several participants (P2, P8, P9, P10) reported that they were unable to distinguish differences in the *Duration* values. In contrast, participants (P3, P5, P6) noted that notifications displayed for 3/90 s felt intrusive, as they seemed to obstruct their FOV. Notably, P1, P3, and P7 expressed concerns about whether notifications displayed for 1/90 s could be reliably noticed. However, P4 remarked, *"I was able to notice the notifications regardless of the display duration, so the shorter duration was not an issue."* These comments highlight individual differences in sensitivity to display duration.

Several participants (P3, P5, P8, P10) reported difficulty distinguishing differences in opacity levels. P2, P6, and P9 observed that high-opacity notifications were disruptive, as they obscured visibility. Conversely, P6 expressed concern that low-opacity notifications might be overlooked. P8 suggested that lower-opacity notifications could help users maintain their focus on primary tasks.

3.6 Summary of Study #1

Participants were able to accurately recognize visual information displayed instantaneously, suggesting that instantaneous visual presentation can be effectively used for notifications. According to the results of Study #1, participants recognized FlashCues even when they were presented the shortest display duration (1/90 s), achieving a recognition error rate of 1.56%. Moreover, the 1/90 s duration resulted in a significantly lower recognition error rate compared to the longer duration condition (2/90 s).

Regarding opacity, participants preferred FlashCue opacity to be set at 75%, while the 25% opacity was less favored due to low visibility. However, even at 25% opacity, the recognition error rate remained below 2.0% when displayed for 1/90 s, indicating that FlashCue can be accurately recognized regardless of opacity level.

4 Study #2

Although FlashCue can present multiple notifications by varying its position and color, it remains unclear whether users can accurately recognize the position and color of FlashCue. In this study, we examine whether participants can recognize the position and color of FlashCue while focusing on other tasks. The study tasks were performed in VR environments.

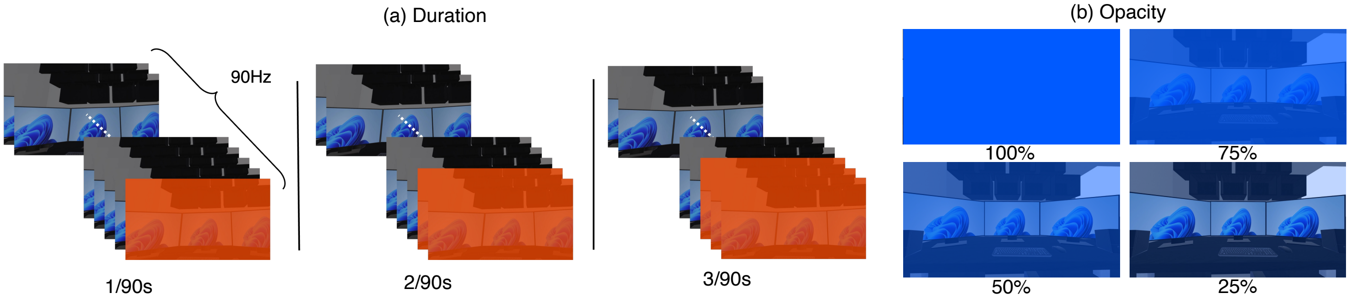

Figure 2: Independent variables of Study #1.

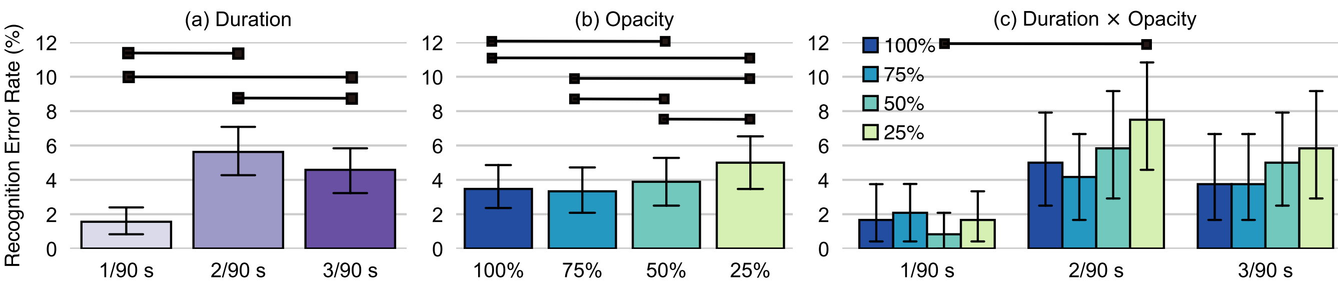

Figure 3: Recognition error rate. Error bars represent standard errors. Statistically significant differences are indicated with solid lines for $p < .05$.

We recruited 12 participants (11 males, one female; mean age = 23.2 years, SD = 1.2 years, ID: P11–P22) from the author's laboratory, none of whom had participated in Study #1. All participants had normal or corrected-to-normal vision. The apparatus used was identical to that in Study#1.

4.1 FlashCue Design

We designed FlashCue display positions using three methods: *2-Split*, *4-Split* (Figure 4a), and *2-and-4-Split*. In the 2-Split method, the screen is divided into two halves, either vertically or horizontally, and FlashCue is displayed in one of these halves, producing four possible display positions. In the 4-Split method, the screen is divided into four quadrants, and FlashCue is displayed in one of these quadrants—also resulting in four possible display positions. In the 2-and-4-Split method, FlashCue is displayed in positions combining the 2-Split and 4-Split methods, yielding eight possible display positions.

In this study, participants were presented with four FlashCue colors (red, blue, yellow, and green) selected from a colorblind-safe palette. Based on feedback from Study #1, in which participants reported that sky blue was difficult to distinguish from blue, sky blue was replaced with green. Referring to the results of Study #1, the display duration of FlashCue was set to 1/90 s, which yielded the lowest recognition error rate, and the opacity was set to 75%, the most preferred opacity value.

4.2 Study Task

The study task involved the performance of a primary task simultaneously with a secondary task. The primary tasks (Figure 4b) were solving a calculation problem (*Calculation* task) or watching a video (*Video-Watching* task). In the Calculation task, participants were instructed to subtract 17 from the currently displayed number repeatedly. The task began with a number approximately equal to 1,000; each time 17 was subtracted from this number, the updated value was displayed. In the Video-Watching task, participants were required to watch a video presented in front of them. Three different TED Talk videos were prepared for each *Method*, and their presentation order was counterbalanced.

The secondary task requires participants to identify the color and position of the FlashCue when it is displayed. Participants were instructed to press a handheld controller's trigger when the FlashCue appeared. After pressing the trigger, a UI was displayed to answer the color and position of the FlashCue that appeared.

4.3 Design

We employed a within-participant design with two independent variables (Fig. 4):

- *Method*: 2-Split, 4-Split, 2-and-4-Split
- *Task*: Calculation, Video-Watching

The order of the *Method* and *Task* was independently counterbalanced. The *Task* order remained consistent across the *Method* conditions for each participant. For the 2-Split and 4-Split methods, the four types of notification positions were presented twice each, while in the 2-and-4-Split method, the eight types were presented

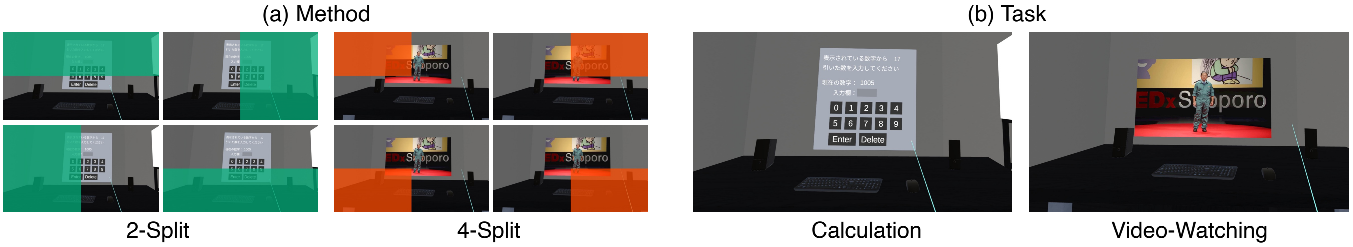

Figure 4: Independent variables of Study #2.

once. The order of the eight trials was randomized. During the eight trials, the four FlashCue colors were each presented twice. The color order was randomized, and the same color was not presented consecutively in the same position for the 2-Split and 4-Split methods. We collected 48 data points per participant (3 *Methods* × 2 *Tasks* × 8 trials), totaling 576 data points across all 12 participants.

4.4 Procedure

The procedure prior to the study task was identical to that in Study #1. Once the task began, the office room of Study #1 was displayed. In the Calculation task, the calculation UI was displayed in front of the participants, while in the Video-Watching task, a video was displayed on a screen positioned in front of them.

Participants performed the primary task corresponding to one of the *Task* conditions while simultaneously completing the secondary task. A FlashCue appeared at a random interval between 25 and 30 seconds after the answer UI disappeared. Participants were instructed to press the trigger on the controller if they recognized the FlashCue. Upon pressing the trigger, the UI for answering the color and position of the FlashCue was displayed. Simultaneously, the calculation UI disappeared in the Calculation task; in the Video-Watching task, the video disappeared. After participants answered the color and position of the FlashCue, the answer UI disappeared, and the calculation UI or video reappeared, depending on the task. The next FlashCue was displayed at a random interval between 25 and 30 seconds after the previous FlashCue. This interval excluded the time taken by the participant to answer.

After completing eight trials under one set of conditions, the participants proceeded to complete eight more trials under a different set of *Task* conditions. Upon finishing all trials for each *Method*, the participants were asked to complete the System Usability Scale (SUS) [1]. A break of at least one minute was provided after each *Method*. This procedure was repeated thrice, corresponding to the three *Methods*. After completing all trials, participants were asked to indicate their preferred *Method* and participated in an interview. Each study took approximately 60 minutes.

4.5 Metrics

We analyzed the following metrics:

- **Recognition Error Rate:** The proportion of trials in which FlashCue was not recognized. These trials are defined as cases in which no response was given within two seconds of its appearance, consistent with the methodology used in Study #1.
- **Color Error Rate:** The proportion of trials in which the color of the FlashCue was incorrectly identified.
- **Position Error Rate:** The proportion of trials in which the position of the FlashCue was incorrectly identified.
- **Overall Error Rate:** The proportion of trials in which an error occurred. An error was defined as a failure to recognize the FlashCue or an incorrect identification of its color or position.
- **Usability Questionnaire:** A subjective measure of usability. This was assessed using the SUS questionnaire for each *Method*. In addition, all participants completed a preference questionnaire to indicate their preferred *Method*.

4.6 Results

We applied the ART method, followed by a two-way repeated measures ANOVA with *Method* and *Task* as factors to analyze recognition error rate, position error rate, color error rate, position error, and overall error rate. Post-hoc comparisons within each factor were conducted using ART-C followed by Holm correction.

We used the Friedman test to evaluate SUS scores. Cross-factor pairwise comparisons were performed using Wilcoxon signed-rank tests, which were also followed by Holm correction.

4.6.1 Recognition Error Rate. The recognition error rates are provided in Table 1. Specifically, the recognition error rates for each *Method*—2-Split, 4-Split, and 2-and-4-Split—were 2.60%, 1.56%, and 1.56%, respectively. We observed the significant effects of *Task* ($F_{1,559} = 544.00, p < .01, \eta_p^2 = .493$) on recognition error rate. Recognition errors were found to be significantly lower during the Calculation task (0.69%) compared to the Video-Watching task (3.13%).

4.6.2 Color Error Rate. The color error rates are provided in Table 1. Specifically, the color error rates for each *Method*—2-Split, 4-Split, and 2-and-4-Split—were 0.52%, 1.56%, and 0.52%, respectively. We observed significant interaction effects of *Method* × *Task* ($F_{2,559} = 29.25, p < .01, \eta_p^2 = .095$) on color error rate; however, post-hoc comparisons did not reveal significant differences.

4.6.3 Position Error Rate. The position error rates are provided in Table 1. Specifically, the position error rates for each *Method*—2-Split, 4-Split, and 2-and-4-Split—were 0.52%, 1.04%, and 4.17%, respectively. We observed significant effects of *Task* ($F_{1,559} = 39.60, p <$

Table 1: Recognition error rate, color error rate, position error rate, and overall error rate for each combination of *Method* and *Task*.

Method	*Task*	*Recognition Error*	*Color Error*	*Position Error*	*Overall Error*
2-Split	Calculation	0.00% (0 / 96)	0.00% (0 / 96)	0.00% (0 / 96)	0.00% (0 / 96)
	Video-Watching	5.21% (5 / 96)	1.04% (1 / 96)	1.04% (1 / 96)	7.29% (7 / 96)
	Total	2.60% (5 / 192)	0.52% (1 / 192)	0.52% (1 / 192)	3.65% (7 / 192)
4-Split	Calculation	1.04% (1 / 96)	1.04% (1 / 96)	2.08% (2 / 96)	4.17% (4 / 96)
	Video-Watching	2.08% (2 / 96)	2.08% (2 / 96)	0.00% (0 / 96)	4.17% (4 / 96)
	Total	1.56% (3 / 192)	1.56% (3 / 192)	1.04% (2 / 192)	4.17% (8 / 192)
2-and-4-Split	Calculation	1.04% (1 / 96)	0.00% (0 / 96)	4.17% (4 / 96)	5.21% (5 / 96)
	Video-Watching	2.08% (2 / 96)	1.04% (1 / 96)	4.17% (4 / 96)	6.25% (6 / 96)
	Total	1.56% (3 / 192)	0.52% (1 / 192)	4.17% (8 / 192)	5.73% (11 / 192)

.01, $\eta_p^2 = .066$) on position error rate. Position errors were significantly higher during the Calculation task (2.08%) compared to the Video-Watching task (1.74%).

4.6.4 Overall Error Rate. The overall error rates are provided in Table 1 and Fig. 5. Specifically, the overall error rates for each *Method*—2-Split, 4-Split, and 2-and-4-Split—were 3.65%, 4.17%, and 5.73%, respectively. We observed the significant effects of *Task* ($F_{1,559} = 28.11, p < .01, \eta_p^2 = .048$) on overall error rate. Overall errors were significantly lower during the Calculation task (3.13%) compared to the Video-Watching task (5.90%).

4.6.5 System Usability Scale (SUS). The overall SUS scores (higher is better) for each *Method*—2-Split, 4-Split, and 2-and-4-Split—were 81.46, 77.71, and 72.30, respectively. A Friedman test was conducted on the overall SUS scores with *Method*, and no significant differences were found across *Method*s ($\chi^2_{2,N=12} = 1.09, p = .58$).

4.6.6 Participants' Preference. The average ranks of preference for each *Method*—2-Split, 4-Split, and 2-and-4-Split—were 2.1 (1st: 4, 2nd: 3, 3rd: 5), 1.8 (1st: 5, 2nd: 5, 3rd: 2), and 2.2 (1st: 3, 2nd: 4, 3rd: 5), respectively. Thus, 4-Split was the most preferred method.

4.6.7 Participants' Feedback. P13, P17, P21, and P22 found the large notifications in the 2-Split method intrusive, while P13, P14, P16, P18, and P22 appreciated the ease of noticing larger notifications. P13 and P14 expressed concerns about the difficulty of noticing smaller notifications in the 4-Split method, while P12, P14, P16, P17, and P19 appreciated the reduced interference with their tasks due to the smaller size of the notifications. P11, P13, P16, P17, and P19 found the variety of notifications in the 2-and-4-Split method overwhelming, while P14, and P18–P21 expressed positive impressions of the increased variety the method provided.

Notably, some participants found it difficult to distinguish between the 2-Split and 4-Split conditions when using the 2-and-4-Split method. For example, P13 remarked, *"I can't really tell the difference between 2-Split and 4-Split. 4-Split seems fine to me,"* while P16 stated, *"Surprisingly, there wasn't much of a difference between 2-Split and 4-Split."*. Additionally, several participants (P11, P13, P17, and P21) observed that notifications displayed on the sides were easier to notice than those displayed above or below. This may be attributed to the wider horizontal FOV of human vision compared to the vertical FOV, as suggested by prior studies on peripheral vision [23].

P20 pointed out that the similarity between the colors of the video in the Video-Watching task and FlashCue's color made it difficult to recognize FlashCue (*"In the video-watching task, the flashing in the video itself seemed to confuse with the notifications."*).

4.7 Summary of Study #2

Participants were able to accurately recognize the color and position of FlashCue, with overall error rates below 6% across all FlashCue methods. These results suggest that FlashCue can reliably support up to 32 types of notifications (4 colors × 8 positions).

Both the recognition error rate and overall error rate were significantly higher during the video-watching task than during the calculation task. In the 2-Split condition, the recognition error rate was 5.21% during the video-watching task and 0.00% during the calculation task. These findings indicate that FlashCue's recognizability is influenced by the nature of the task and background context.

5 Discussion

5.1 Recognition Rate of Momentary Visual Information

The results of Study #1 and Study #2 demonstrate that FlashCue, which delivers momentary visual information, can be recognized with high accuracy. In Study #1, the participants recognized FlashCue displayed for only 1/90 s (≈ 0.011 s) at 50% opacity at a low recognition error rate of 1.23%. In Study #2, the participants recognized the 2-Split condition, which included four colors and four display positions, at an error rate of 0% while performing tasks involving calculation. Additionally, the participants recognized the 2-and-4-Split condition, which included the four colors and eight position variations, with an error rate of 5.73%. These findings indicate that FlashCue can deliver highly recognizable notifications even with an extremely short display duration. Moreover, the results indicate FlashCue's potential for conveying even larger amounts of information.

FlashCue's recognition error rate, however, may increase due to several factors, such as the user's blinking. Since a blink typically lasts about 0.1 seconds, users may fail to recognize FlashCue if it is displayed during a blink. To address this issue, FlashCue should be presented multiple times for a single notification. Another factor contributing to recognition errors is the similarity between the

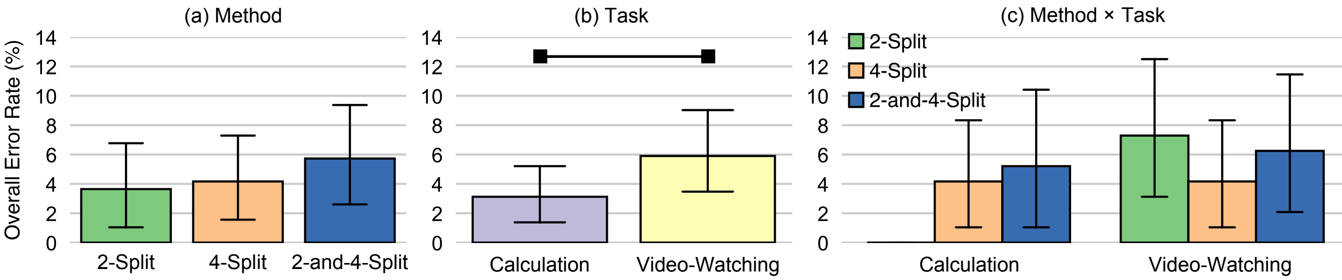

Figure 5: Overall error rate. Error bars represent standard errors. Statistically significant differences are indicated with solid lines for $p < .05$.

FlashCue color and the background. In Study #2, the color error rate of FlashCue was significantly higher during the Video-Watching task compared to the Calculation task. This increase can be attributed to confusion between the colors of the video content and those of FlashCue. This aligns with prior findings that humans' recognition of brief visual stimuli is affected by background content and color characteristics [34].

5.2 Strengths and Weaknesses of FlashCue

FlashCue offers several advantages over other notification methods in terms of both delivery speed and the number of distinguishable notification types. Based on the results of Study #1 and Study #2, FlashCue with a display duration of 1/90 seconds delivered information accurately and much faster than typical audio or vibration notifications. In addition, FlashCue can display up to 32 distinct notification types—a capacity that exceeds many existing approaches. Prior research has shown that increasing the number of vibration-based notifications often reduces recognition accuracy, with more than eight types generally not recommended [43]. In contrast, FlashCue uses simple visual parameters, such as position and color, to allow users to quickly and reliably distinguish among a large set of notifications. Overall, the combination of extremely short notification times and support for a wide range of notification types are distinctive strengths of FlashCue compared over other notification methods.

On the other hand, FlashCue is not appropriate for all contexts and has certain drawbacks compared with other notification methods. First, it can only convey simple visual information, such as color and position, and cannot effectively present content that requires time to understand, such as text. Second, FlashCue is less suitable for delivering critical notifications compared to other notification systems, as FlashCues may be missed during blinking and do not strongly capture the user's attention. Consequently, FlashCue should not be used for urgent information such as phone calls or disaster alerts, but rather for less critical notifications that can be overlooked without serious consequences, such as application notifications or incoming emails.

5.3 Limitations and Future Work

This study has several limitations. First, the participant sample and experimental apparatus may limit the generalizability of the work's findings. Most participants were young male university students, resulting in a narrow range of gender and age demographics. Because the proportion of instantaneous visual information that can be recognized may vary with age and visual acuity, further research is needed to examine differences in FlashCue recognition performance across diverse populations.

Second, this study did not directly compare FlashCue with other notification methods. Although FlashCue may not outperform all alternative notification systems, it can still demonstrate relatively strong performance in suitable contexts. Therefore, future work should compare the performance of FlashCue with other notification methods across a range of tasks and identify the contexts in which FlashCue is most effective.

Finally, although FlashCue has a variety of design parameters, this study investigated only a subset of these parameters. Our research, which serves as an initial exploration of momentary visual displays for notifications, focused solely on two parameters: display position and color. However, FlashCue supports a broader range of design possibilities, including multiple sequential displays, the presentation of more complex visual information, and dynamic positioning based on the user's gaze. Future research should explore these design dimensions and advance notification methods based on instantaneous visual cues.

6 Conclusion

In this study, we proposed FlashCue, a notification method that conveys simple visual information through momentary presentation. We examined suitable parameters for FlashCue's use, including display duration, opacity, and the number of distinguishable notifications. The results indicated that FlashCue can deliver up to 32 distinct notifications with high recognition accuracy. This work serves as an initial step toward exploring the use of instantaneous visual information for notifications. Future research should further investigate FlashCue's applications, evaluate its performance in diverse contexts, and identify its limitations.

Acknowledgments

This work was partially supported by JSPS KAKENHI Grant Number 24K21319.

References

[1] John Brooke. 1995. SUS: A quick and dirty usability scale. *Usability Evaluation in Industry, CRC Press,* 189 (1995), 189–194.

[2] Kuan-Wen Chen, Yung-Ju Chang, and Liwei Chan. 2022. Predicting Opportune Moments to Deliver Notifications in Virtual Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 186, 18 pages. doi:10.1145/3491102.3517529

[3] Marina Cidota, Stephan Lukosch, Dragos Datcu, and Heide Lukosch. 2016. Workspace Awareness in Collaborative AR using HMDs: A User Study Comparing Audio and Visual Notifications. In *Proceedings of the 7th Augmented Human International Conference 2016 (AH '16)*. Association for Computing Machinery, New York, NY, USA, 3:1–3:8. doi:10.1145/2875194.2875204

[4] Enrico Costanza, Samuel A. Inverso, Elan Pavlov, Rebecca Allen, and Pattie Maes. 2006. eye-q: Eyeglass Peripheral Display for Subtle Intimate Notifications. In *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services* (Helsinki, Finland) *(MobileHCI '06)*. Association for Computing Machinery, New York, NY, USA, 211–218. doi:10.1145/1152215.1152261

[5] Edward Cutrell, Mary Czerwinski, and Eric Horvitz. 2001. Notification, Disruption, and Memory: Effects of Messaging Interruptions on Memory and Performance. In *IFIP TC13 International Conference on Human-Computer Interaction*. https://api.semanticscholar.org/CorpusID:17478824

[6] Shakiba Davari, Feiyu Lu, and Doug A. Bowman. 2020. Occlusion Management Techniques for Everyday Glanceable AR Interfaces. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 324–330. doi:10.1109/VRW50115.2020.00072

[7] Lisa A. Elkin, Matthew Kay, James J. Higgins, and Jacob O. Wobbrock. 2021. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21)*. Association for Computing Machinery, New York, NY, USA, 754–768. doi:10.1145/3472749.3474784

[8] Barrett Ens and Pourang Irani. 2017. Spatial Analytic Interfaces: Spatial User Interfaces for In Situ Visual Analytics. *IEEE Computer Graphics and Applications* 37, 2 (2017), 66–79. doi:10.1109/MCG.2016.38

[9] Joseph L. Gabbard, Mark Smith, and John E. Swan. 2018. Effects of frequent focal distance switches on human performance in optical see-through augmented reality. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2018)*. IEEE, Munich, Germany, 1–10. doi:10.1109/ISMAR.2018.00034

[10] Joseph L. Gabbard, John E. Swan, and David Hix. 2004. The effects of text drawing styles, background textures, and natural lighting on text legibility in outdoor augmented reality. *Presence: Teleoperators and Virtual Environments* 13, 4 (2004), 372–386. doi:10.1162/1054746041944849

[11] Joseph L. Gabbard, J. Edward Swan II, and Deborah Hix. 2005. The Effects of Text Drawing Styles, Background Textures, and Natural Lighting on Text Legibility in Outdoor Augmented Reality. In *Proceedings of IEEE VR*. IEEE, 11–18. doi:10.1109/VR.2005.1492748

[12] Sarthak Ghosh, Lauren Winston, Nishant Panchal, Philippe Kimura-Thollander, Jeff Hotnog, Douglas Cheong, Gabriel Reyes, and Gregory D. Abowd. 2018. NotifiVR: Exploring Interruptions and Notifications in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics, IEEE Educational Activities Department,* 24, 4 (2018), 1447–1456. doi:10.1109/TVCG.2018.2793698

[13] Michelle R. Greene and Aude Oliva. 2009. The Briefest of Glances: The Time Course of Natural Scene Understanding. *Psychological Science,* 20, 4 (2009), 464–472. doi:10.1111/j.1467-9280.2009.02316.x PMID: 19399976.

[14] Kalanit Grill-Spector and N Kanwisher. 2005. Visual Recognition – As Soon as You Know It Is There, You Know What It Is. *Psychological Science, Association for Psychological Science,* 16, 2 (2005), 152–160. doi:10.1111/j.0956-7976.2005.00796.x

[15] Jan Gugenheimer, Evgeny Stemasov, Harpreet Sareen, and Enrico Rukzio. 2018. FaceDisplay: Towards Asymmetric Multi-User Interaction for Nomadic Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3173574.3173628

[16] James J. Higgins and Suleiman Tashtoush. 1994. An Aligned Rank Transform Test for Interaction. *Nonlinear World,* 1, 2 (1994), 201–211.

[17] Joyce Ho and Stephen S. Intille. 2005. Using Context–Aware Computing to Reduce the Perceived Burden of Interruptions from Mobile Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, Oregon, USA) *(CHI '05)*. Association for Computing Machinery, New York, NY, USA, 909–918. doi:10.1145/1054972.1055100

[18] Sture Holm. 1979. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics, John Wiley & Sons,* 6, 2 (1979), 65–70.

[19] Shamsi T. Iqbal and Brian P. Bailey. 2011. Oasis: A Framework for Linking Notification Delivery to the Perceptual Structure of Goal–Directed Tasks. *ACM Transactions on Computer-Human Interaction, Association for Computing Machinery, New York, NY, USA,* 17, 4 (2011), 15:1–15:28. doi:10.1145/1879831.1879833

[20] Shamsi T. Iqbal and Eric Horvitz. 2010. Notifications and Awareness: A Field Study of Alert Usage and Preferences. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* (Savannah, Georgia, USA) *(CSCW '10)*. Association for Computing Machinery, New York, NY, USA, 27–30. doi:10.1145/1718918.1718926

[21] Kei Ito. 2012. Color Universal Design - Towards a Color-Vision-Barrier-Free Society. *Journal of Information Processing and Management* 55, 5 (2012), 307–317. doi:10.1241/johokanri.55.307

[22] Nuwan Janaka, Chloe Haigh, Hyeongcheol Kim, Shan Zhang, and Shengdong Zhao. 2022. Paracentral and near-peripheral visualizations: Towards attention–maintaining secondary information presentation on OHMDs during in–person social interactions. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 551:1–551:14. doi:10.1145/3491102.3502127

[23] J. Adam Jones, J. Edward Swan II, and Mark Bolas. 2013. Peripheral stimulation and its effect on perceived spatial scale in virtual environments. *IEEE Transactions on Visualization and Computer Graphics, Institute of Electrical and Electronics Engineers,* 19, 4 (2013), 701–710.

[24] J. Bern Jordan and Gregg C. Vanderheiden. 2024. International Guidelines for Photosensitive Epilepsy: Gap Analysis and Recommendations. *ACM Trans. Access. Comput.* 17, 3, Article 17 (Oct. 2024), 35 pages. doi:10.1145/3694790

[25] Olivier R. Joubert, Guillaume A. Rousselet, Denis Fize, and Michele Fabre-Thorpe. 2007. Processing scene context: Fast categorization and object interference. *Vision Research, Elsevier B.V.,* 47, 26 (2007), 3286–3297. doi:10.1016/j.visres.2007.09.013

[26] Stephen Kaplan. 1992. Environmental Preference in a Knowledge–Seeking, Knowledge-Using Organism. In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford University Press, New York, NY, USA, 581–598.

[27] Jonathan W. Kelly, Lucia A. Cherep, Alex F. Lim, Taylor Doty, and Stephen B. Gilbert. 2021. Who Are Virtual Reality Headset Owners? A Survey and Comparison of Headset Owners and Non-Owners. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. 687–694. doi:10.1109/VR50410.2021.00095

[28] Thomas Kosch, Andrii Matviienko, Florian Müller, Jessica Bersch, Christopher Katins, Dominik Schön, and Max Mühlhäuser. 2022. NotiBike: Assessing Target Selection Techniques for Cyclist Notifications in Augmented Reality. 6, MHCI, Article 197 (2022), 24 pages. doi:10.1145/3546732

[29] Kostadin Kushlev, Jason Proulx, and Elizabeth W. Dunn. 2016. "Silence Your Phones": Smartphone Notifications Increase Inattention and Hyperactivity Symptoms. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1011–1020. doi:10.1145/2858036.2858359

[30] Renzo C. Lanfranco, Andrés Canales-Johnson, Hugh Rabagliati, Axel Cleeremans, and David Carmel. 2024. Minimal exposure durations reveal visual processing priorities for different stimulus attributes. *Nature Communications, Nature Research,* 15, 1 (2024), 8523:1–8523:13. doi:10.1038/s41467-024-52778-5

[31] Hyunjin Lee and Woontack Woo. 2023. Exploring the Effects of Augmented Reality Notification Type and Placement in AR HMD while Walking. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. 519–529. doi:10.1109/VR55154.2023.00067

[32] Zhipeng Li, Yi Fei Cheng, Yukang Yan, and David Lindlbauer. 2024. Predicting the Noticeability of Dynamic Virtual Elements in Virtual Reality. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, 954:1–954:17. doi:10.1145/3613904.3642399

[33] Philipp Müller, Sander Staal, Mihai Bâce, and Andreas Bulling. 2022. Designing for Noticeability: Understanding the Impact of Visual Importance on Desktop Notifications. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, 472:1–472:13. doi:10.1145/3491102.3501954

[34] Aude Oliva and Philippe G. Schyns. 2000. Diagnostic Colors Mediate Scene Recognition. *Cognitive Psychology, Elsevier B.V.,* 41, 2 (2000), 176–210. doi:10.1006/cogp.1999.0728

[35] Martin Pielot, Karen Church, and Rodrigo de Oliveira. 2014. An In–Situ Study of Mobile Phone Notifications. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services* (Toronto, ON, Canada). Association for Computing Machinery, New York, NY, USA, 233–242. doi:10.1145/2628363.2628364

[36] Martin Pielot, Amalia Vradi, and Souneil Park. 2018. Dismissed!: A Detailed Exploration of How Mobile Phone Users Handle Push Notifications. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Barcelona, Spain) *(MobileHCI '18)*. Association for Computing Machinery, New York, NY, USA, 3:1–3:11. doi:10.1145/3229434.3229445

[37] Leon Pietschmann, Michel Schimpf, Zhu-Tian Chen, Hanspeter Pfister, and Thomas Bohné. 2025. Enhancing User Performance and Human Factors through Visual Guidance in AR Assembly Tasks *(CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, Article 226, 8 pages. doi:10.1145/3706599.3720094

[38] Lucas Plabst, Florian Niebling, Sebastian Oberdörfer, and Francisco Ortega. 2025. Order Up! Multimodal Interaction Techniques for Notifications in Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 31, 5 (2025), 2258–2267. doi:10.1109/TVCG.2025.3549186

[39] Lucas Plabst, Sebastian Oberdörfer, Francisco Raul Ortega, and Florian Niebling. 2022. Push the Red Button: Comparing Notification Placement with Augmented and Non-Augmented Tasks in AR. In *Proceedings of the 2022 ACM Symposium on Spatial User Interaction* (Online, CA, USA) *(SUI '22)*. Association for Computing

Machinery, New York, NY, USA, Article 17, 11 pages. doi:10.1145/3565970.3567701

[40] Mary C. Potter, Brad Wyble, Carl Erick Hagmann, and Emily S. McCourt. 2014. Detecting meaning in RSVP at 13 ms per picture. *Attention, Perception, & Psychophysics, the Psychonomic Society,* 76, 2 (2014), 270–279. doi:10.3758/s13414-013-0605-z

[41] Rufat Rzayev, Susanne Korbely, Milena Maul, Alina Schark, Valentin Schwind, and Niels Henze. 2020. Effects of Position and Alignment of Notifications on AR Glasses during Social Interaction. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (Tallinn, Estonia) *(NordiCHI '20)*. Association for Computing Machinery, New York, NY, USA, 30:1–30:11. doi:10.1145/3419249.3420095

[42] Rufat Rzayev, Sven Mayer, Christian Krauter, and Niels Henze. 2019. Notification in VR: The Effect of Notification Placement, Task and Environment. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Barcelona, Spain) *(CHI PLAY '19)*. Association for Computing Machinery, New York, NY, USA, 199–211. doi:10.1145/3311350.3347190

[43] Bahador Saket, Chrisnawan Prasojo, Yongfeng Huang, and Shengdong Zhao. 2013. Designing an effective vibration-based notification interface for mobile phones. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (San Antonio, Texas, USA) *(CSCW '13)*. Association for Computing Machinery, New York, NY, USA, 149–1504. doi:10.1145/2441776.2441946

[44] K. C. Salter and R. F. Fawcett. 1993. The ART test of interaction: a robust and powerful rank test of interaction in factorial models. *Communications in Statistics – Simulation and Computation, Taylor & Francis,* 22, 1 (1993), 137–153. doi:10.1080/03610919308813085

[45] Hyunjoo Shin and Jaewan Park. 2024. Evaluation of Interface Design for Video See-Through Head-Mounted Display: Focusing on Walking Users. *Archives of Design Research* 37, 1 (2024), 181–191. doi:10.15187/adr.2024.02.37.1.181

[46] Sunghyun Song, Geeyoung Noh, Junwoo Yoo, Ian Oakley, Jundong Cho, and Andrea Bianchi. 2015. Hot & tight: exploring thermo and squeeze cues recognition on wrist wearables. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers (ISWC '15)*. Association for Computing Machinery, New York, NY, USA, 39–42. doi:10.1145/2802083.2802092

[47] Dan Tasse, Anupriya Ankolekar, and Joshua Hailpern. 2016. Getting Users' Attention in Web Apps in Likable, Minimally Annoying Ways. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 3324–3334. doi:10.1145/2858036.2858174

[48] Simon J. Thorpe, Denis Fize, and Catherine Marlot. 1996. Speed of Processing in the Human Visual System. *Nature, Nature Research,* 381 (1996), 520–522. https://api.semanticscholar.org/CorpusID:4303570

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