Beyond the Beep: Comparing Visual-Only and Multimodal Notification Design for Improved Noticeability in AR

Aditya Raikwar [*]
Colorado State University

Zahra Borhani [†]
Colorado State University

Lucas Plabst [‡]
Colorado State University

Anil Ufuk Batmaz [§]
Concordia University

Mayra Donaji Barrera Machuca [¶]
University of Calgary

Florian Niebling [‖]
University of Applied Sciences

Francisco Raul Ortega [**]
Colorado State University

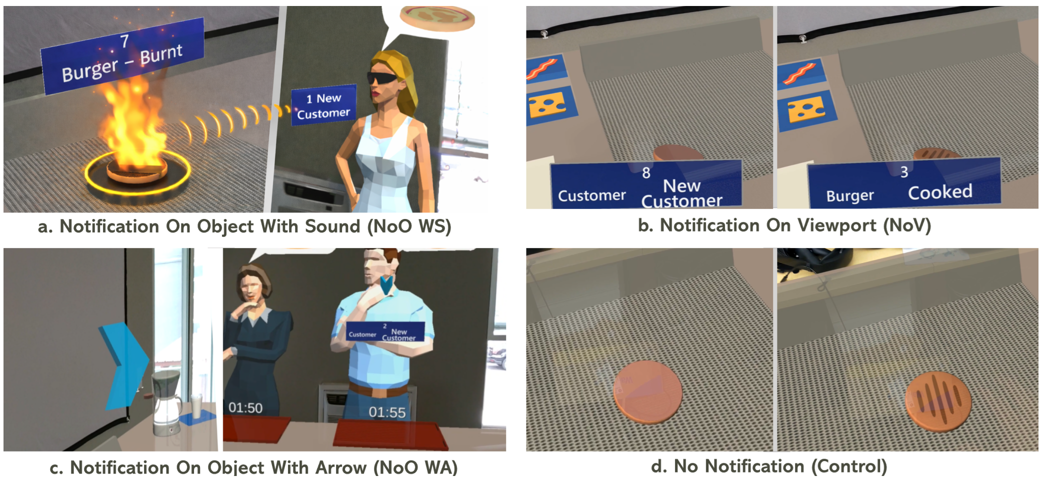

Figure 1: Four notification conditions: a) Notification on Object with Sound (NoO WS) b) Notification on Viewport (NoV) c) Notification on Object with Arrow (NoO WA) d) The Control condition (No Notification)

Abstract

Augmented Reality (AR) notifications must balance attracting user attention with the need to avoid distraction or cognitive overload, particularly in performance-critical tasks. While prior work suggests that multimodal notifications combining visual and audio cues can improve user attention, the effectiveness of visual-only notification designs remains less understood. This study investigates whether visual-only notifications can support user performance and attention as effectively as multimodal notifications in AR environments. We conducted a within-subject user study to compare four notification conditions: Notifications on Object With Sound (NoO WS), Notification on Viewport (NoV), Notification on Object With Arrow (NoO WA), and a Control condition (No Notifications). The study was implemented in ARtisan Bistro, an open-source AR sandbox environment simulating restaurant tasks. Results indicate that user performance did not differ significantly across notification conditions, suggesting that multimodal designs do not necessarily outperform visual-only notifications in all scenarios. In addition, combined visual-audio notifications did not significantly outperform NoV in terms of attention capture or user preference. According to the findings, despite their practicality, not all visual-only notifications can function as well as multimodal ones. The study emphasizes the need for more research into increasingly difficult AR activities, improvements in visual notification design, and the effects of background noise on notification effectiveness.

Index Terms: Augmented Reality, Human-computer Interaction (HCI), Visualization design and evaluation methods, Notification.

[*]e-mail: adityaraikwar@outlook.com
[†]e-mail: zahra.borhani@colostate.edu
[‡]lucas.plabst@colostate.edu
[§]ufuk.batmaz@concordia.ca
[¶]mbarrera@ucalgary.ca
[‖]florian.niebling@th-koeln.de
[**]F.Ortega@colostate.edu

1 Introduction

Augmented Reality is rapidly evolving into a critical technology across various domains, enhancing user experiences by overlaying digital information onto the real world. Despite these advancements, effectively delivering information in AR environments remains a significant challenge, particularly in the design of notifications [33, 37, 45, 31, 5]. For instance, notifications must provide valuable, relevant information to users while effectively conveying the intended message or alert, ensuring users understand its importance or the required action [34]. However, while notifications need to capture attention, they should not be so intrusive or frequent that they disrupt the user's current task or workflow. Excessive or poorly designed notifications can divert the user's focus from their primary task, leading to decreased productivity or errors [30]. Thus, notifications must balance being informative and attention-grabbing without causing undue distraction or cognitive overload [34]. Maintaining this balance is essential in environments that require high-performance and situational awareness, such as customer service and industrial applications.

In AR environments, notifications can be delivered through multiple modalities, including visual [29], audio [34], or haptic feedback [27]. Modern head-mounted displays (HMDs) support these modalities beyond visual cues by providing built-in earphones for audio and vibration motors in hand-held controllers for haptic feedback. By leveraging multiple channels, AR systems can more effectively guide user attention. For instance, multimodal notifications that combine visual and auditory cues have been shown to enhance user focus and reduce task load when interacting with HMDs [24, 18]. However, the effectiveness of visual-only notifications, especially in environments where auditory cues may be impractical due to ambient noise or other constraints, remains largely underexplored [20, 3]. Yet, in many real-world environments, such as industrial settings, outdoor locations, or crowded public spaces, ambient noise can have detrimental effects on the users' attention [4]. This raises a critical design challenge: understanding the tangible costs and benefits of relying solely on visual notifications when audio is removed.

While previous work by Raikwar et al. [34] demonstrated that multimodal layouts combining spatial audio and object-anchored visuals represent the gold standard for AR notifications, it did not investigate the attentional and cognitive costs of removing auditory cues entirely. Furthermore, prior research failed to explore visual-only fallback mechanisms (e.g., periphery arrow-guided cues) to resolve the "out-of-view" notification problem when sound is unavailable. By isolating these constraints, this study directly addresses gaps left unexamined in the literature to chart an optimal design space for high-noise AR environments. To address this gap, we adopt the same high-fidelity ARtisan Bistro simulation as a benchmark testbed but extend their scope by investigating the trade-offs that occur when sound is eliminated in favor of alternative visual notification designs. Ultimately, this study aims to pose and answer the central research question: *What is the cost of silence?*

To achieve this, we compare the pre-established multimodal design (NoO WS) with two alternative visual-only patterns: a common non-contextual Notification on Viewport (NoV) and a Notification on Object With an Arrow (NoO WA), designed to solve the "out-of-view" problem without using sound (Figure 1). We conducted a user study with 24 participants to measure the effects of these designs on task completion, attention capture (noticeability and reaction time), and subjective user experience (usability ratings and preferences). The study was conducted inside a simulated AR environment called ARtisan Bistro developed by Raikwar et al. [35][1].

Results from this experiment demonstrate that gross task performance under the conditions tested was not significantly affected, but the choice of notification modality has significant implications for attention and usability. We found that the multimodal (NoO WS) and viewport-fixed (NoV) designs yielded significantly improved noticeability and reaction times compared to the arrow-guided (NoO WA) design. Moreover, users strongly preferred the multimodal system, as they rated it higher for usability. These results reveal a key trade-off: while some visual-only designs may be tolerable, others actively degrade the user experience by increasing cognitive load and slowing response times.

1.1 Research Questions

We aimed to address the following research questions (RQs) to gain insights that can improve user experience (UX) within AR notification environments:

RQ_1 : **How does notification modality (multimodal vs. two distinct visual-only designs) affect user task performance (measured by customers served) in a multitasking AR environment?** Prior research suggests that visual and auditory notifications can improve user focus and reduce task load [24, 18, 34]. This research question explores the effectiveness of visual-only notifications in supporting user performance and satisfaction within an AR environment. Specifically, when presented solely as a visual modality, can these notifications affect users' ability to manage tasks and interruptions while serving customers?

[1]https://github.com/NuiLab/ARtisan-Bistro

RQ_2 : **What are the effects of different notification designs on attention capture, measured by notification observation rate and reaction time?** In scenarios where audio may be unavailable or distracting, such as very loud environments or social settings, visual notifications become the primary means of grabbing users' attention. This research question determines if visual-only notification types (NoV, NoO WA) are as noticeable and effective as notifications that combine visual cues with sounds (NoO WS).

RQ_3 : **How does notification modality affect a user's subjective perception of the notification?** This perception could influence how readily users adopt and feel comfortable with notifications in AR environments over the long run. For example, some users might find visual-only notifications less intrusive, while others prefer the combined audio-visual approach for enhanced clarity. The perceived effectiveness and user acceptance of notification systems are significantly impacted by user preferences in previous studies [27, 34]

By addressing these questions, we hope to better understand the trade-offs between pure visual notifications and notifications that combine visual and auditory cues. This knowledge will inform the design of AR notification systems that optimize user experience. The ideal system should support task management effectively while minimizing distractions. It should also promote user comfort and acceptance by considering preferences for notification modality and intrusiveness. The key contributions of this research include:

1. Comparative Analysis of Notification Designs: Through a systematic comparison of a "gold standard" multimodal design to two distinct visual-only designs, we illustrate the trade-offs (in attention, reaction time, and usability) of removing the audio from AR notifications.

2. Attention Capture and Usability Outcomes: The findings reveal that viewport-fixed notifications can continue to maintain high noticeability without sound. Periphery arrow cues, on the other hand, can significantly increase reaction times and reduce usability, illustrating the cognitive cost of ambiguous visual cues in the absence of sound.

2 Related Work

This section reviews prior research on notification design in immersive environments.

2.1 Visual Cues

Traditionally, visual notifications have been used to convey information on flat displays, such as smartphones, tablets, and monitors. These notifications range from simple indicators, such as a bell icon or a simple ping with no animation, to more detailed messages, like phone notifications that display lengthy text [1, 2]. While such notifications are effective on personal devices, their design becomes more challenging in shared-display environments, where concerns such as privacy must be considered. Weber et al. [43] conducted focused-group studies examining user perception of TV notifications with respect to their frequency, placement, duration, and information content. Their findings suggest that notifications should be limited to essential information, presented during appropriate breaks, and designed to preserve user privacy in shared settings.

As visual notifications have expanded to wearable and mixed reality systems, researchers have investigated how critical information can be delivered without disrupting users' primary tasks. Wallmyr et al. [41] found that head-up mixed reality displays reduced workload and improved responsiveness compared to traditional head-down displays. In addition, notification placement also plays a critical role in effectiveness. Harrison et al. [14] found that the location of visual cues can affect user performance, with cues attached to the wrist producing the fastest reaction times compared to cues placed on the arm, brooch, shoulder, thigh, waist, and shoe, respectively. Their findings further suggest that factors such as Visual accessibility, occlusion, and physical distance to the cues influence the effectiveness of visual notifications [14].

2.1.1 Information Placement

A recurring question in AR HMD research is not only *what* information to show, but *where* to place it so that it can be comfortably perceived without disrupting ongoing activity. Several previous studies have shown that bottom-aligned notifications can significantly improve noticeability and comprehension [23, 38, 17]. Rzayev et al. [38] found that text placement in AR HMDs affects reading performance, with top-right positioning increasing workload and reducing comprehension, compared to the bottom-center, center, and bottom-right locations. Similarly, Imamov et al. [17] investigated 18 world-lock positions to determine optimal placement for glanceable information. They discovered that the middle and central-low positions were the fastest and favored, whereas those in the top-left were the slowest. Participants also felt more at ease gazing down at the interface from a distance of two meters rather than three meters. However, since this study focused on simple activities, further investigation is required to determine how location affects more complex tasks.

Plabst et al. [30] investigated notification placement in AR HMDs during sustained concentration tasks. They discovered that the most favored and successful placements for notifications were bottom-center or in relation to the real world. Their task, however, was fully stationary, so findings about real-world notification placement might not apply to mobile tasks, where optimal placement cannot be guaranteed. Similarly, Lee et al. [25] discovered that during AR walking exercises, bottom field of view (FoV) placement increased noticeability and understanding for icon- and text-based messages. Extending this to real-world mobility, a recent study in outdoor cycling scenario showed that while bottom-aligned notifications improve perceived safety and noticeability compared to top-aligned notifications, AR HMD alerts can still induce attentional tunneling regardless of placement [16].

2.2 Audio Cues

While there are studies that have shown that users often prefer visual notifications over audio ones [10], audio notifications can provide an alternative when visual notifications might be distracting or unsafe, such as while driving or operating heavy machinery. Lee et al. [27] found that graded alerts improved safety and reduced annoyance compared to single-stage alerts. Audio cues have been present to help people who are visually impaired. Crommentuijn et al. [11] conducted a study to evaluate several approaches for giving haptic or audible indications when there is an obstruction in the path. The findings suggest that well-designed audio cues can significantly improve visually impaired individuals' ability to navigate their environment.

While audio cues are effective on their own, recent work increasingly emphasizes the benefits of multimodal notification systems [22, 34, 27]. Raikwar et al. [34] demonstrated that combining visual and auditory cues improved user performance and was preferred over unimodal alternatives in AR environments. Similarly, Cho et al. investigated the impact of audio and visual notifications in extended reality (XR) and showed that adding audio cues can significantly enhance noticeability and reduce response times for out-of-view visual notifications [9].

2.3 Notifications in Virtual and Augmented Reality

Virtual Reality (VR) environments completely immerse users in a virtual world. Considering this, notifications in VR cannot use the same framework as other smart devices, as it can break immersion. There have been studies proposing different ways to display notifications in VR. Zenner et al. proposed a framework in which, instead of stopping the application, the message is incorporated into the scenario [44]. In a medieval scenario, the information is delivered to the user by a villager who passes them a letter; in a futuristic one, the user receives a notice from a drone that flies to them. They [44] stated that the presentation of alerts depends on both the urgency of the notice and the context of the virtual world. While this preserves the user's immersion, it also adds more work for developers. The developers need to incorporate notification systems that are highly context-dependent and based on user preferences. This also removes standardization and best practices for notification design and efficient techniques.

To draw users' attention to events occurring outside virtual spaces, Ghosh et al.'s NotifiVR system extends alerts to include visual, audio, and haptic cues [13]. Ghosh et al. suggested three approaches for audio cues: Skeuomorphism, which uses metaphorical designs, such as the sound of a door opening to indicate someone entering; Gradual Intensification, which uses sounds that gradually intensify to indicate urgency; and Spatial Sound, which plays sound from the location of the notification in three dimensions. Their recommendations for visual cues included 3D Representation, where notifications are displayed in 3D space instead of being locked to the viewport; Skeumorphism, such as showing a clock face for an alarm; 3D Pop-ups, placing notifications on virtual surfaces like walls; and Flash and Pause, which halts the user's immersion during dangerous situations, allowing them to safely reorient to the real world.

There are not enough studies focusing on notification display in AR settings. In the NotifEye research by Lucero et al., participants tracked alerts that popped up in their range of vision while walking along a busy road [28]. Participants interacted with a rubbing pad that enabled basic functions, including opening and closing alerts. They discovered that a non-distracting notification system can be created by integrating the rub pad with basic information. Participants did, however, have reservations over the social acceptability of AR gadgets, particularly concerning their size, comfort, and style.

A study by Raikwar et al. [34] investigated the impact of notification modality and placement on user performance and preferences in an AR HMD environment. They compared six conditions that varied notification location (on-object, docked, or absent) and modality (visual-only or visual with audio). Their results showed that notifications placed directly on objects and accompanied by audio cues significantly improved task performance and were preferred by users. These results highlighted the benefits of spatially contextual and multimodal notifications for improving noticeability and maintaining user attention in AR. Building on this research, we further examined the effectiveness of visual-only and multimodal notifications, providing a more thorough assessment of their influence on user performance and attention.

3 Notification Design

This section dives into the design of the notification system implemented in this study. To directly investigate the "cost of silence,", as our multimodal baseline, we used the best-performing and most preferred notification design from Raikwar et al. [34]: Notification on Object With Sound. We then designed two distinct visual-only notifications to test specific paradigms against this baseline.

All notifications were displayed in the experiment as a text billboard with a uniform dark blue background and white text, in line with previous text legibility studies [12, 19, 6]. The notification updated the content and reset the dismissal timer if a new notification from the same source arrived before the old one vanished. Visual notifications remained active for 7 seconds.

3.1 Notification on Object with Sound (NoO WS): Context-Aware Notifications

This is the established "gold-standard" multimodal baseline from Raikwar et al.'s work [34]. It displays a visual notification in the 3D space directly above the related object (Figure 1.a), in combination with a localized audio signal. A single, localized sound clip for notifications was employed to avoid overcomplication and shorten the learning curve in the Notifications on Object experiment. This approach enabled participants to leverage spatial audio cues and determine the direction of the sound source. The HoloLens 2 uses a head-related transfer function to simulate spatial audio with stereo speakers, which has been found to closely match real-world localized sound sources in the horizontal direction and in depth perception [36]. This combination of a contextually placed visual cue and a non-intrusive sound is designed to draw the user's attention effectively without requiring them to shift their line of sight away from their workspace. The chosen notification sound was a subtle bubble-popping sound to minimize distraction and maintain neutrality.

3.2 Notification on Viewport

This condition represents a non-contextual notification option, akin to a Heads-Up Display (HUD). The notification is anchored to a position in the user's FoV, in this instance, the bottom-center (Figure 1.b), based on recommendations from previous research [26, 30]. This position ensures the notification is noticeable, but perhaps at the cost of blocking the user's line of sight to the primary task or interfering with context. It offers a valuable point of comparison for a standard visual-only implementation approach.

3.3 Notification on Object with Arrow

Warden et al. examined how different cue designs and placement affect visual search tasks in AR [42]. Their study compared three cue types (arrow, icon image, minimap) and two placement locations (center, downward) within an AR HMD. They found that arrows provided the most significant search benefit with high accuracy and short search time. Adding more information to the display can lead to obstructing primary tasks, causing information overload, and distracting attention from primary tasks or essential environmental cues. This condition is a direct attempt to solve the "out-of-view" problem of contextual notifications without using audio. It uses the same on-object positioning as the baseline, but adds a guiding arrow at the periphery of the user's FoV when the notification source is out of view (Figure 1.c). Arrows have been shown to be effective for visual search tasks [42], so this design tests the viability of using a directional visual cue as an alternative to the attention-directing properties of spatialized sound. This is an important test of a specific visual-only paradigm developed to aid in target acquisition.

4 Methodology

This section describes the study methodology used to compare the four notification conditions.

4.1 Experiment Design

This study employed a within-subjects design, where the independent variable was the notification type. The experimental conditions were chosen to facilitate a systematic comparison of notification modalities with prior research and to investigate the impact of audio across different visual designs. This resulted in four experimental conditions: (1) Notification on Object With Sound (NoO WS) – Participants received visual notifications displayed on the object, accompanied by an audio notification. (2) Notification on Viewport (NoV) – Visual notifications were displayed in a fixed location at the bottom center of the participant's FoV. (3) Notification on Object With Arrow (NoO WS) – Participants received visual notifications displayed directly on the virtual object requiring attention. An arrow was paired with each notification, guiding the participants to the respective notification. (4) No Notifications (Control) – This baseline condition presented no visual notifications. Participants completed the tasks solely by monitoring changes within the environment.

4.2 Procedure

The experiment followed these steps: (1) Participants arrived at the research laboratory, provided informed consent, and received a detailed explanation of the study.(2) Participants watched a short video introducing the AR environment and the different notification types they would encounter throughout the experiment. (3) Participants completed a practice session using a HoloLens 2 headset. This session involved preparing food items for three virtual customers without a time limit, allowing them to learn how to use the controls, and ask questions about the task. (4) Following the training session, participants completed the four experimental conditions. The order of conditions was counterbalanced across participants using a Latin-square design. Before each condition, participants received instructions describing the corresponding notification design. Each condition presented a simulated customer service scenario in which six virtual customers requested meals. Meals had to be prepared within 120 seconds before the customer left; otherwise, they were counted as not successfully served. Participants used virtual ingredients to prepare the food while receiving notifications about cooking status and new customers. After each condition, participants completed two surveys: System Usability Scale (SUS) [7] and NASA Task Load Index (NASA-TLX) [15]. (5) Upon completing all four conditions, participants answered post-experiment interview questions to convey their subjective experience with the different notification designs.

Prior to the main study, a pilot study with three participants was conducted to validate the task design, refine the experimental procedure, and ensure the logging system was functioning correctly. To assess user performance and experience, the measures and logs recorded during the experiment are listed in Appendix A flowchart depicting the steps has been provided in Section 7.

4.3 Participants

The experiment involved 24 participants, with 13 identifying as male and 11 female. No participant identified as non-binary or any other gender. Their age ranged from 19 to 56 years old ($M = 27.88, SD = 8.99$). Participants were recruited from various sources: students and staff from [REDACTED] and individuals not affiliated with the university. The majority of participants were either students or faculty (91.67%).

Prior experience with AR and VR technology was assessed to gauge the participants' familiarity with immersive environments. Over 90% of the participants (91.67%) reported having used at least one AR or VR device in the past. This indicates some level of comfort and a basic understanding of how to interact within these environments. The participants were asked about previous work experience in a restaurant setting (e.g., as a cook or waiter) to understand their potential familiarity with the task context (restaurant service). Approximately one-third of the participants (33.3%) reported having some form of restaurant work experience. While not a prerequisite for successful task completion, this information provided context for interpreting any potential influence of prior experience on user performance. Each participant received a compensation of $20. The university's internal review board approved this study.

4.4 Apparatus

The experiment was conducted in the ARtisan Bistro environment, an open-source AR sandbox designed to simulate complex multitasking scenarios. The use of the experimental environment was a deliberate decision for our task validity. We chose ARtisan Bistro because it is an established, increasingly common benchmark for complex, multitasking AR studies [34, 32, 29]. The environment is used as a metaphor for various real-world sustained-attention tasks, commonly found in manufacturing or health care, among others. Users have to pay attention to several time-critical processes and frequently transition from stationary to mobile.

Furthermore, the features of the ARtisan Bistro task are well-suited to examine the efficacy of notifications specifically. The surroundings are designed to place a user in a scenario in which they must divide their attention between two or more simultaneous events, e.g., preparing different foods and serving customers. The exercise requires participants to physically wander through the virtual setting, so essential events frequently occur at stations positioned outside the user's immediate line of sight. This combination of divided attention and off-screen activity exactly defines the circumstances in which a notification system's ability to capture attention and guide a user is most relevant, allowing for a rigorous test of the difference between a multimodal cue, a viewport-fixed cue, and a visual-only cue in terms of pointing attention in the direction of an out-of-view target.

The AR environment was presented through a HoloLens 2 AR-HMD. Participants used a virtual hand or a raycast to interact with virtual objects. The HoloLens 2 uses a head-related transfer function to simulate spatial audio with stereo speakers, while research indicates challenges with elevation perception due to speaker placement, users are generally capable of localizing sounds along the horizontal plane [36]. Since the on-object notifications were confined to the XZ plane (horizontal and depth), the HoloLens 2 speakers were adequate for this specific scenario.

5 Results

This section presents the effects of the different notification techniques on both objective measures (user performance, reaction time, and notification observation rate), and subjective measures (usability, taskload, and user preference). Table 1 summarizes the collected measures across the four experimental conditions.

We performed one-way ANOVA to investigate the main effects. To control assumptions, we tested for normal distribution using Shapiro Wilk test [39]. Levene's test [8] was conducted to assess the homogeneity of variances. In case of violations of assumptions, a non-parametric test, the Kruskal-Wallis H test [21] was used.

5.1 Objective Measures

5.1.1 Performance

The analysis focused on the number of customers served (out of six) successfully (see Figure 2 and Table 1) as a measure of user performance.

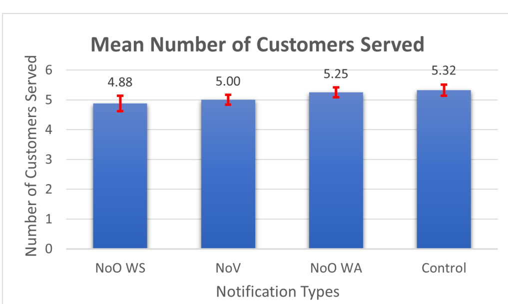

Figure 2: Amount of customers served. Bars represent the standard error of the mean.

The analysis focused on the number of customers successfully served (out of a maximum of six) as a measure of gross user performance. Initial assumption checks revealed severe violations of normality across all conditions, which is expected for bounded, discrete count data. Consequently, a non-parametric within-subjects omnibus test was utilized. A Friedman test of differences among repeated measures was conducted. The results indicated no statistically significant difference in the number of customers served across the four notification conditions, $\chi^2(3) = 1.77,\ p = 0.62$. Descriptive statistics confirmed that mean performance was practically identical regardless of the notification modality (NoO WS: $M = 4.87, SD = 1.29$; NoV: $M = 5.00, SD = 0.83$; NoO WA: $M = 5.25, SD = 0.84$; Control: $M = 4.87, SD = 1.72$).

To address potential concerns regarding statistical power, a sensitivity analysis indicated that our sample size of $N = 24$ provides 80% power to detect medium within-subject effect sizes ($\alpha = 0.05$). Paired post-hoc testing revealed that the observed within-subjects effect sizes between conditions were exceptionally small, ranging from nonexistent to small (Control vs. NoO WS: $d = 0.00$; Control vs. NoV: $d = -0.09$; NoV vs. NoO WS: $d = -0.11$; Control vs. NoO WA: $d = -0.27$; NoV vs. NoO WA: $d = 0.29$; NoO WA vs. NoO WS: $d = 0.34$). The lack of significant differences is therefore driven by the exceptionally small magnitude of the observed effects, suggesting that, in this specific simulated environment, the choice of notification modality did not materially affect users' ability to serve customers.

We further investigated whether prior restaurant experience influenced user performance. A t-test was performed to compare the effect of prior restaurant work experience on the number of customers served. The t-test revealed no statistically significant difference in the number of customers served between the two groups ($t(22) = 0.42, p = 0.68$). This suggests that prior experience in the culinary field did not significantly affect the user's ability to complete customer service tasks in the simulated environment.

5.1.2 Notifications Observed

The analysis focused on the number of times participants actively acknowledged notifications (see Figure 3 and Table 1) by verbally calling out the word "notification" every time they observed a new notification (this can be a visual or auditory observation). The experimenter logged the acknowledgment using a keyboard key, which recorded the time from when the notification appeared to when it was seen.

A one-way ANOVA was conducted to assess potential differences in notification observation rates across the three experimental conditions (NoO WS, NoV, and NoO WA). The ANOVA results revealed a statistically significant main effect of notification type on the percentage of notifications observed ($F(2,56) = 4.32, p < 0.05$). This suggests that at least one of the notification types differed from the others in terms of how many notifications participants noticed.

Further analysis using Tukey's Honestly Significant Difference (HSD) test for multiple comparisons identified a significant difference in notification observation rates between the NoO WS and NoO WA conditions ($p < 0.05, 95\%\ CI = [1.70, 19.79]$). This indicates that participants observed a significantly higher percentage of notifications when presented with both visual cues on the object and accompanying sound cues (NoO WS) compared to the condition with only visual cues on the object and an arrow (NoO WA).

Interestingly, the ANOVA results did not reveal any statistically significant differences in notification observation rates between the NoO WS and NoV conditions ($p = 0.124$), nor between NoV and NoO WA ($p = 0.567$). This suggests that participants observed a similar percentage of notifications when presented with visual and audio cues on the object (NoO WS) compared to only visual cues displayed at a fixed position in the user's FoV (NoV) and with vi-

Table 1: Mean performance, notifications observed, reaction times, SUS scores, and workload scores for different notification techniques. Values are Mean (SD: Standard Deviation, SE: Standard Error)

Notification Techniques	Customers	Call-outs	Reaction Time (s)	SUS Score	TLX Score
NoO WS	4.88 (*1.27*, *0.26*)	81.63% (*18.90*, *4.03*)	2.34 (*0.70*, *0.15*)	80.00 (*15.28*, *3.12*)	53.02 (*18.53*, *3.78*)
NoV	5.00 (*0.82*, *0.17*)	76.78% (*15.03*, *3.13*)	2.64 (*0.88*, *0.18*)	77.50 (*14.49*, *2.96*)	52.29 (*16.82*, *3.43*)
NoO WA	5.25 (*0.83*, *0.17*)	68.53% (*20.26*, *4.53*)	3.30 (*0.96*, *0.21*)	66.67 (*18.34*, *3.74*)	54.97 (*17.69*, *3.61*)
Control	5.32 (*0.87*, *0.18*)	- (-, -)	- (-, -)	75.00 (*17.14*, *3.50*)	53.19 (*18.80*, *3.84*)

sual cues on the object with an arrow (NoO WA) compared to visual cues displayed at a fixed position in the user's FoV (NoV).

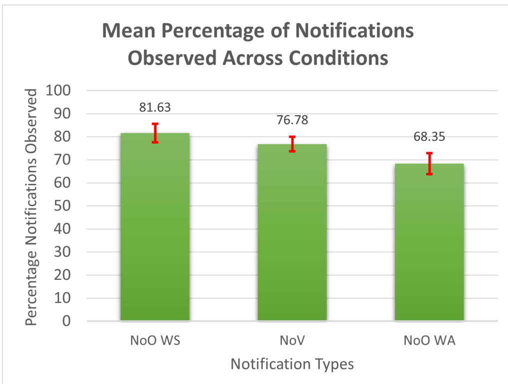

Figure 3: Average number of notifications observed. Bars represent the standard error of means.

5.1.3 Notifications Observed Reaction Time

The analysis focused on the time participants took to verbally report the confirmation number displayed in each visual notification. The average reaction time across the different notification design conditions is illustrated in Figure 4. Detailed information on individual means, standard deviations, and standard errors can be found in Table 1. A one-way ANOVA was conducted to assess potential differences in reaction times (time taken to notice a new notification) across the three experimental conditions (NoO WS, NoV, and NoO WA). The ANOVA results revealed a statistically significant main effect of notification type on reaction time for observed notifications ($F(2,58) = 7.13, p < 0.01$). This indicates that at least one notification type resulted in significantly faster or slower reaction times than the others. Further analysis using Tukey's HSD test for multiple comparisons identified significant differences in notification response time between the following conditions:

NoO WS vs NoO WA: Participants responded significantly faster to notifications presented with both visual cues on the object and accompanying sound cues (NoO WS) compared to the condition with only visual cues on the object and an arrow (NoO WA) ($p < 0.01, 95\%\ CI = [-1.195, -0.225]$).

NoV vs NoO WA: Participants also responded significantly faster to visual cues displayed at a fixed position in the user's FoV (NoV) compared to the condition with visual cues on the object and an arrow (NoO WA) ($p < 0.01, 95\%\ CI = [-1.113, -0.133]$).

Interestingly, the ANOVA results did not reveal any statistically significant difference in notification response time between the NoO WS and NoV conditions ($p = 0.896$). This suggests that participants responded with similar speeds to notifications presented with both sound and visual cues on the object (NoO WS) compared to only visual cues displayed at a fixed position in the user's FoV (NoV).

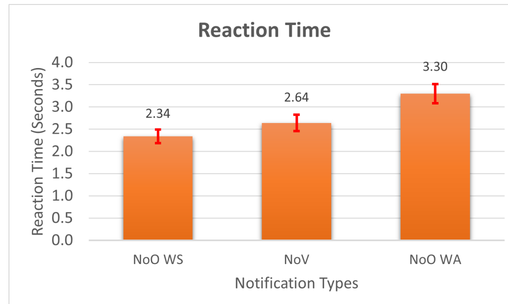

Figure 4: Average reaction time to register the notification in seconds. Bars represent the standard error of the mean.

5.2 Subjective Measures

5.2.1 System Usability

The system's (notifications) usability was evaluated using the System Usability Scale (SUS) questionnaire [7]. Participants completed the 10-item SUS survey after each condition to record their subjective experience. Participants rated the system favorably on the SUS. The generally accepted rating is $> 80.3 = Excellent$; $68 - 80.3 = Good$; $68 = Average$; $51 - 68 = Poor$; $< 51 = Awful$ [40]. The results showed variations in perceived usability across conditions (see Table 1). The complete SUS survey is available in the supplemental materials.

A one-way ANOVA was conducted to assess potential differences in SUS scores across the four experimental conditions (NoO WS, NoV, NoO WA, and Control). The ANOVA results revealed a statistically significant main effect of notification type on the SUS scores ($F(3,92) = 2.87, p < 0.05$). This suggests that at least one of the notification types differed from the others in terms of SUS score.

Further analysis using Tukey's HSD test for multiple comparisons identified a significant difference in SUS scores between the NoO WS and NoO WA conditions ($p < 0.05, 95\%\ CI = [0.694, 25.972]$). This indicates that participants experienced better system usability when presented with both visual cues on the object and accompanying sound cues (NoO WS) compared to the condition with only visual cues on the object and an arrow (NoO WA). Interestingly, the ANOVA results did not reveal any statistically significant differences in SUS scores between the NoO WS and NoV conditions ($p = 0.955$), the NoO WS and Control ($p = 0.729$), the NoV and NoO WA conditions ($p = 0.120$), the NoV and Control ($p = 0.955$), and the NoO WA and Control ($p = 0.317$).

5.2.2 Taskload

The NASA-TLX questionnaire was used to assess participants' perceived workload during system interactions (notifications). After completing each condition, participants provided ratings on six subscales (see Table 1). Each subscale ranged from 0.5 to 10 with intervals of 0.5. The overall workload was calculated by using Raw NASA-TLX [15]. A one-way ANOVA was conducted to assess potential differences in NASA-TLX scores across the four experimental conditions (NoO WS, NoV, NoO WA, and Control). The ANOVA results did not reveal a statistically significant main effect of notification type on the NASA-TLX scores ($F(3,92) = 0.092, p = 0.97$).

5.2.3 Preference

To gain insights into user preferences for notification design, participants were asked to rank the notification conditions from "Best" to "Worst". The results revealed a clear order of preference. Participants overwhelmingly favored the NoO WS condition, followed by NoV, Control, and lastly, NoO WA. The ranked order of preference for each notification type is shown in Table 2.

Table 2: Ranked order of preference for each notification type.

Notification Techniques	1^{st} rank	2^{nd} rank	3^{rd} rank	4^{th} rank
NoO WS	10	5	6	3
NoV	8	5	7	4
NoO WA	5	9	4	6
Control	1	5	7	11

6 Discussion

Our study investigated the costs and benefits of removing sounds from AR notifications by comparing a "gold-standard" multimodal version to two alternative visual-only versions. In scenarios where audio may not be viable, notifications may potentially rely solely on visual channels.

6.1 Performance

Building on the established value of notification cues for user experience, RQ_1 examined the effectiveness of visual-only notifications on user performance in AR environments. While past research suggests benefits of both visual and auditory notifications for focus and task load reduction [24, 18, 34], this study explored whether visual-only notifications could maintain comparable performance. Maintaining, or even surpassing, performance with vision-only notifications is important in environments where the sound modality may be impractical.

The analysis of the number of customers served revealed no significant difference between the three notification conditions (NoO WS, NoV, NoO WA). This suggests that user performance in serving customers remained comparable whether they received visual-only notifications or sound-and-visual notifications in the simulated restaurant environment. These findings indicate that visual-only notifications may not hinder user performance compared with multimodal notification designs. We cannot conclude that the notification types are "equally effective".

However, it is important to acknowledge that these results also indicate no significant difference in performance between conditions in which participants received notifications and those in which they did not (Control condition). This lack of distinction between notification and control groups is unexpected and warrants further investigation. Notably, these findings diverge from the results of previous studies [24, 18].

One potential explanation for these observations is the cognitive demand for customer service tasks within the AR environment. The tasks may not have been complex enough for visual-only notifications to be a significant disadvantage relative to sound-based notifications. This could also explain the lack of difference between participants who received notifications and those who did not. More complex tasks might benefit more from the combined support of visual and auditory cues.

While no significant differences in gross performance were detected across conditions, post-hoc chronological analysis revealed severe ceiling and learning effects that heavily influenced these outcomes. Across all conditions, 41.67% of trials achieved a perfect score, severely restricting variance. Furthermore, a chronological evaluation of trials independent of the experimental condition revealed a significant procedural learning effect (Friedman $\chi^2 = 9.79, p = 0.020$), with users performing significantly better by their third chronological trial ($M = 5.37$) compared to their first ($M = 4.58$).

This indicates that the simulated task was sufficiently simple procedurally for users to achieve mastery and adapt to suboptimal notification conditions (such as the Control or NoO WA). Consequently, we conclude that in highly learnable AR tasks, gross performance metrics are insensitive to the nuances of notification design. Instead, the true impact of notification modality manifests in the cognitive and temporal costs required to maintain that performance, which are captured in our Noticeability and Usability metrics.

6.2 Notifications Observed

RQ_2 focused on the effectiveness of visual-only notifications (NoV, NoO WA) in capturing user attention. The analysis of notification observation rates revealed that participants acknowledged a significantly higher percentage of notifications when presented with the combined audio and visual cues (NoO WS) compared to the condition with only an arrow guiding to the visual notifications (NoO WA), suggesting that the choice of visual cues has a high cost in terms of both noticeability and speed. This aligns with previous research suggesting the benefits of multimodal notifications for enhanced user awareness [24, 18]. Interestingly, the observation rates for FoV-fixed visual notifications (NoV) did not differ significantly from those in the combined modality condition (NoO WS). Even though it does not suggest that both designs have an equivalent performance, the absence of a measured deficit in such attention-capture metrics for NoV is noteworthy.

However, it is important to consider potential limitations. The tasks in this study might not require enough attention to yield a clear advantage for the notifications. More demanding tasks might show a greater difference. Building on these findings, future research could explore how visual notification design can be optimized to improve attention capture in AR environments, for example, by experimenting with visual design elements such as size, color, animation, and location within the AR space to maximize their attention-grabbing potential. Another avenue is to explore the effectiveness of NoV notifications across various noise environments and varying task complexities.

Participants responded significantly faster to both NoV and NoO WS notifications compared to NoO WA. One reason for slow response time may be that arrows are located at the periphery. If the user was too focused on the primary task, they might have missed the appearance of a new arrow. Another reason might be information overload: the additional visual element (the arrow) may have caused a momentary overload, especially compared to the simpler cues in NoV and NoO WS. It replaced a low-effort auditory alert with a high-effort visual task.

The observation rates for purely visual notifications (NoV) did not differ significantly from those in the combined-modality condition (NoO WS). It is noteworthy that the NoV design showed no measurable disadvantage in capturing attention, even if this doesn't confirm equivalent performance between the two designs. This means that a viewport-fixed, well-positioned alert could be an acceptable fallback when audio cues are not feasible. It successfully captured user attention without the statistically significant increase in reaction time observed in the other visual-only design. Therefore, in environments where audio might prove ineffective, the NoV design is a great compromise, presenting a way to transmit information that can be sensed without utilizing the auditory channel.

This research question highlights a trade-off between notification design and user experience. While visual-only notifications can be effective in some situations, adding auditory cues can enhance attention capture. Future research can further refine the design of visual notifications to optimize the user experience across diverse AR contexts.

6.3 User Preference, Usability, and Workload

The effects of various notification design strategies on user perception were assessed using usability, workload, and user preference metrics to address RQ_3. These metrics provided crucial new information about how participants saw and interacted with the various notification situations. Participants rated the overall system favorably, as indicated by the SUS scores (except NoO WA). However, these scores varied across notification conditions. The highest SUS score was observed with the NoO WS condition. This finding aligns with the user-reported preference ranking, where NoO WS was the most favored condition.

A one-way ANOVA confirmed a statistically significant main effect of notification type on SUS scores. Further analysis revealed that the NoO WS condition resulted in significantly higher perceived usability than the NoO WA condition. The SUS scores for the NoV and Control conditions were not statistically different from the NoO WS condition. The NoO WA design seemed to introduce a new secondary task by having the user follow the arrow, rather than helping with the primary task.

Participants described the arrows as "overwhelming" and noted that the system collapsed when multiple notifications appeared, stating, "I don't even know where you are pointing anymore". The significantly slower reaction time provides a direct measurement of the time cost associated with this added cognitive processing. However, the added load is not reflected in the TLX scores. Additionally, no other pair of conditions showed a statistically significant difference in their SUS scores. This could be because the SUS is not sensitive enough to detect subtle usability differences, or because a learning curve could have masked initial usability differences. It is, however, worth noting that scores above 68 are considered "good usability", which only NoO WA fails to reach ($M = 66.67$).

The user-reported preference rankings mirrored the usability findings. Participants overwhelmingly favored the NoO WS condition, followed by NoV, Control, and NoO WA. This alignment suggests that users found the most usable notification design (a combined visual and auditory one) to be preferable. These findings highlight the potential benefits of considering both visual and auditory cues when designing notification systems for AR environments, particularly for tasks where high usability and user satisfaction are crucial. While both usability and user preference favored the combined visual and auditory cues (NoO WS), even simpler notification designs (NoV and Control) achieved comparable task completion. Perceived workload remained similar across conditions.

6.4 Interviews

Participants were interviewed in follow-up sessions to better understand users' rationale and specific design aspects that impact user experience. A deductive method for analyzing qualitative interviews was used to structure and analyze the interview. All the questions asked during the interview are provided in Section 7.

The interviews corroborated the survey results. Participants overwhelmingly favored the NoO WS condition due to its perceived clarity and efficiency. They explained that the combination of visual notification on the object and the accompanying sound subtly drew their attention away from the task at hand, minimizing the distraction from the primary task. Participant 8 (P8) stated, *"I'll be focusing on some object, but when it is a sound, I will sense that - Okay, there's a new notification."* P16 echoed this sentiment, saying, *"It let me know what's going on behind the scenes without intruding on what I was doing at the moment."*

In contrast, participants found the NoO WA condition the least favorable, reporting that the arrow pointing to the object was distracting and momentarily blocked their view. This aligns with the finding that NoO WA had a statistically significantly slower response time than NoV. P4 commented, *"The arrows felt a bit overwhelming. It felt like it was telling you to do a lot of stuff."* P12 stated, *"I think they (arrows) were too big, but I think the main thing is when they piled on top of one another, I'm like, I don't even know where you are pointing anymore."*

The interviews revealed a divide among participants regarding the NoV visual notifications. Some participants were able to completely ignore or work around these notifications. P22 said, *"I will just bend down so that the notification goes down, and then I will do the task. And it went away after 7 seconds."*. However, others found themselves easily distracted. P17 stated, *"It's (NoV notification) obstructing me from doing other things."*

Participants also offered valuable suggestions for improving the AR notification system. They indicated that notifications could be made more dynamic, such as by incorporating pulsating/ gradual change effects (P2, P16) or different color/metaphorical sounds to enhance their ability to capture attention (P6, P17, P18, P26), especially in visually cluttered environments. Additionally, participants expressed a desire for customization options (P8, P19), including the ability to adjust notification frequency (P2), select which stations trigger notifications (P15), and set notification priorities based on their individual preferences and specific work contexts (P14).

7 Conclusion

This study compared four different notification designs (Notifications on Object With Sound, Notification on Viewport, Notification On Object With Arrow, and No Notifications) to examine their effects on information delivery efficiency, distraction levels, and user satisfaction. Results showed that user performance remained consistent across all notification types. This suggests that visual-only notifications do not hinder user performance compared to multimodal notifications (RQ_1). However, the lack of a significant difference between the notification and no notification conditions highlights the potential need for more cognitively demanding tasks to reveal the performance benefits of notifications.

Regarding attention capture (RQ_2), participants acknowledged a significantly higher percentage of notifications with combined audio and visual cues (NoO WS) than visual-only notifications guided by an arrow (NoO WA). Interestingly, fixed-FoV visual notifications (NoV) showed attention rates similar to those of both NoO WS and NoO WA, suggesting their potential effectiveness in noisy environments where auditory cues may be less effective. Response times were faster for NoV and NoO WS than for NoO WA, indicating that more straightforward visual notifications, or those with audio support, are more effective at quickly capturing user attention. As measured by SUS scores, user preference, and usability, the NoO WS condition was favored, which also received the highest user preference ranking (RQ_3). The NASA-TLX scores showed no significant differences in perceived workload across notification types, suggesting that the tasks were not sufficiently demanding to differentiate workload impacts.

Supplemental Materials

All supplemental materials are available for reference. These include (1) the image of the play area, (2) the complete SUS survey and interview questions, (3) the experiment flowchart, and (4) a short demo video.

Acknowledgments

This work was supported by the National Science Foundation (NSF) awards NSF 2439474, NSF 2444132, NSF 2327569 NSF 2238313, and NSF 2223432.

References

[1] Android and Google. Android notifications. https://developer.android.com/develop/ui/views/notifications. 2

[2] Apple. Apple notifications. https://support.apple.com/en-us/108781. 2

[3] G. F. Arfvidsson, M. Ljungdahl Eriksson, H. Lidbo, and K. Falkenberg. Design considerations for short alerts and notification sounds in a retail environment. In *18th Sound and Music Computing Conference, SMC 29 June-1 July 2021*, vol. 2021, pp. 261–267. Sound and Music Computing Network, 2021. 2

[4] K. Azizkhani. *Assessing the effects of external environmental factors in public spaces on mental health and well being*. PhD thesis, Politecnico di Torino, 2022. 2

[5] Z. Borhani, A. Ebrahimpour-Boroojeny, and F. R. Ortega. Small cues, big differences: Evaluating interaction and presentation for annotation retrieval in ar. *arXiv preprint arXiv:2509.11401*, 2025. 1

[6] Z. Borhani, P. Sharma, and F. R. Ortega. Survey of annotations in extended reality systems. *IEEE Transactions on Visualization and Computer Graphics*, 30(8):5074–5096, 2023. 3

[7] J. Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996. 4, 6

[8] M. B. Brown and A. B. Forsythe. Robust tests for the equality of variances. *Journal of the American Statistical Association*, 69(346):364–367, 1974. doi: 10.1080/01621459.1974.10482955 5

[9] H. Cho, D. Edgar, D. Lindlbauer, and J. O'Hagan. Evaluating dynamic delivery of audio+ visual message notifications in xr. In *2025 IEEE conference virtual reality and 3D user interfaces (VR)*, pp. 277–287. IEEE, 2025. 3

[10] M. Cidota, S. Lukosch, D. Datcu, and H. Lukosch. 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*, pp. 1–8, 2016. 3

[11] K. Crommentuijn and F. Winberg. Designing auditory displays to facilitate object localization in virtual haptic 3d environments. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility*, Assets '06, p. 255–256. Association for Computing Machinery, New York, NY, USA, 2006. doi: 10.1145/1168987.1169042 3

[12] J. L. Gabbard, J. E. Swan, and D. Hix. The Effects of Text Drawing Styles, Background Textures, and Natural Lighting on Text Legibility in Outdoor Augmented Reality. *Presence: Teleoperators and Virtual Environments*, 15(1):16–32, Feb. 2006. doi: 10.1162/pres.2006.15.1.16 3

[13] S. Ghosh, L. Winston, N. Panchal, P. Kimura-Thollander, J. Hotnog, D. Cheong, G. Reyes, and G. D. Abowd. Notifivr: exploring interruptions and notifications in virtual reality. *IEEE transactions on visualization and computer graphics*, 24(4):1447–1456, 2018. 3

[14] C. Harrison, B. Y. Lim, A. Shick, and S. E. Hudson. Where to locate wearable displays? reaction time performance of visual alerts from tip to toe. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 941–944, 2009. 2, 3

[15] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, 2006. doi: 10.1177/154193120605000909 4, 6

[16] L. He, M. Siegenthaler, L. Y. Ho, L. Liu, E. Bosch, T. Kosch, B. Ens, S. Goodwin, B. Tag, and D. S. Elvitigala. Designing augmented reality for cyclists: How text-based notification placements influence attentional tunneling and cycling experience. *Proceedings of the ACM on Human-Computer Interaction*, 9(8):292–312, 2025. 3

[17] S. Imamov, D. Monzel, and W. S. Lages. Where to display? how interface position affects comfort and task switching time on glanceable interfaces. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 851–858. IEEE, 2020. 3

[18] S. T. Iqbal and E. Horvitz. Notifications and awareness: a field study of alert usage and preferences. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, pp. 27–30, 2010. 2, 7

[19] J. Jankowski, K. Samp, I. Irzynska, M. Jozwowicz, and S. Decker. Integrating text with video and 3d graphics: The effects of text drawing styles on text readability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, p. 1321–1330. Association for Computing Machinery, New York, NY, USA, 2010. doi: 10.1145/1753326.1753524 3

[20] R. Jung. Non-intrusive audio notification in emotion classified background music. *Proceedings of Meetings on Acoustics*, 9(1):050001, 04 2010. doi: 10.1121/1.3430888 2

[21] W. H. Kruskal and W. A. Wallis. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260):583–621, 1952. doi: 10.1080/01621459.1952.10483441 5

[22] M. J. Lazaro, S. Kim, J. Lee, J. Chun, and M.-H. Yun. Interaction Modalities for Notification Signals in Augmented Reality. In *Proceedings of the 2021 International Conference on Multimodal Interaction*, pp. 470–477. ACM, Montréal QC Canada, Oct. 2021. doi: 10.1145/3462244.3479898 3

[23] H. Lee, S. Bang, and W. Woo. Effects of coordinate system and position of ar notification while walking. *Virtual Reality*, 27(2):829–848, 2023. 3

[24] H. Lee and W. Woo. Investigating display position of a head-fixed augmented reality notification for dual-task. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 584–585. IEEE, 2022. 2, 7

[25] H. Lee and W. Woo. 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)*, pp. 519–529, Mar. 2023. doi: 10.1109/VR55154.2023.00067 3

[26] H. Lee and W. Woo. 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)*, pp. 519–529, 2023. doi: 10.1109/VR55154.2023.00067 4

[27] J. D. Lee, J. D. Hoffman, and E. Hayes. Collision warning design to mitigate driver distraction. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems*, pp. 65–72, 2004. 1, 2, 3

[28] A. Lucero and A. Vetek. Notifeye: using interactive glasses to deal with notifications while walking in public. In *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology*, pp. 1–10, 2014. 3

[29] L. Plabst, F. Niebling, S. Oberdörfer, and F. Ortega. Order up! multimodal interaction techniques for notifications in augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 31(5):2258–2267, 2025. doi: 10.1109/TVCG.2025.3549186 1, 5

[30] L. Plabst, S. Oberdörfer, F. R. Ortega, and F. Niebling. 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*, pp. 1–11. ACM, Online CA USA, Dec. 2022. doi: 10.1145/3565970.3567701 1, 3, 4

[31] L. Plabst, L. Plabst, F. Niebling, and F. R. Ortega. Five-day research-in-the-wild observation of notifications on smartglasses: A double edged sword. In *Proceedings of the 2025 31st ACM Symposium on Virtual Reality Software and Technology*, VRST '25. Association for Computing Machinery, New York, NY, USA, 2025. doi: 10.1145/3756884.3765976 1

[32] L. Plabst, A. Raikwar, S. Oberdörfer, F. R. Ortega, and F. Niebling. Exploring unimodal notification interaction and display methods in augmented reality. In *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*, VRST '23. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3611659.3615683 5

[33] L. S. Plabst, F. R. Ortega, F. Niebling, M. Arefin, B. Trinkenreich, H. Soto, et al. Investigating notification presentation and interaction in augmented reality. 2025. 1

[34] A. Raikwar, L. Plabst, A. U. Batmaz, F. Niebling, and F. R. Ortega. Ping! your food is ready: Comparing different notification techniques in 3d ar cooking environment, 2024. 1, 2, 3, 4, 5, 7

[35] A. Raikwar, L. Plabst, and F. R. Ortega. Artisan bistro: A cooking task environment to conduct studies in augmented reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 909–910. IEEE, 2022. 2

[36] W. Ryu, S. Lee, and E. Park. The Effect of Training on Localizing HoloLens-Generated 3D Sound Sources. *Sensors (Basel, Switzerland)*, 24(11):3442, May 2024. doi: 10.3390/s24113442 4, 5

[37] R. Rzayev, S. Mayer, C. Krauter, and N. Henze. Notification in vr: The effect of notification placement, task and environment. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pp. 199–211, 2019. 1

[38] R. Rzayev, P. W. Woźniak, T. Dingler, and N. Henze. Reading on smart glasses: The effect of text position, presentation type and walking. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–9, 2018. 3

[39] S. S. SHAPIRO and M. B. WILK. An analysis of variance test for normality (complete samples)†. *Biometrika*, 52(3-4):591–611, 12 1965. doi: 10.1093/biomet/52.3-4.591 5

[40] W. T. Interpreting system usability scale (sus) score. https://uiuxtrend.com/measuring-system-usability-scale-sus/. 6

[41] M. Wallmyr, T. A. Sitompul, T. Holstein, and R. Lindell. Evaluating mixed reality notifications to support excavator operator awareness. In *IFIP Conference on Human-Computer Interaction*, pp. 743–762. Springer, 2019. 2

[42] A. C. Warden, C. D. Wickens, D. Mifsud, S. Ourada, B. A. Clegg, and F. R. Ortega. Visual search in augmented reality: Effect of target cue type and location. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 66(1):373–377, 2022. doi: 10.1177/1071181322661260 4

[43] D. Weber, S. Mayer, A. Voit, R. Ventura Fierro, and N. Henze. Design guidelines for notifications on smart tvs. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*, pp. 13–24, 2016. 2

[44] A. Zenner, M. Speicher, S. Klingner, D. Degraen, F. Daiber, and A. Krüger. Immersive notification framework: adaptive & plausible notifications in virtual reality. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–6, 2018. 3

[45] J. Zheng, X. Wang, S. Mayer, and L.-H. Lee. Non-urgent messages do not jump into my headset suddenly! adaptive notification design in mixed reality. *arXiv preprint arXiv:2603.05893*, 2026. 1