Push the Red Button: Comparing Notification Placement with Augmented and Non-Augmented Tasks in AR

Lucas Plabst
HCI Group, University of Würzburg
Würzburg, Germany
lucas.plabst@uni-wuerzburg.de

Sebastian Oberdörfer
HCI Group, University of Würzburg
Würzburg, Germany
sebastian.oberdoerfer@uni-wuerzburg.de

Francisco Ortega
Computer Science & NUILAB, Colorado State University
Fort Collins, CO, USA
fortega@colostate.edu

Florian Niebling
University of Applied Sciences Fulda
Fulda, Germany
florian.niebling@informatik.hs-fulda.de

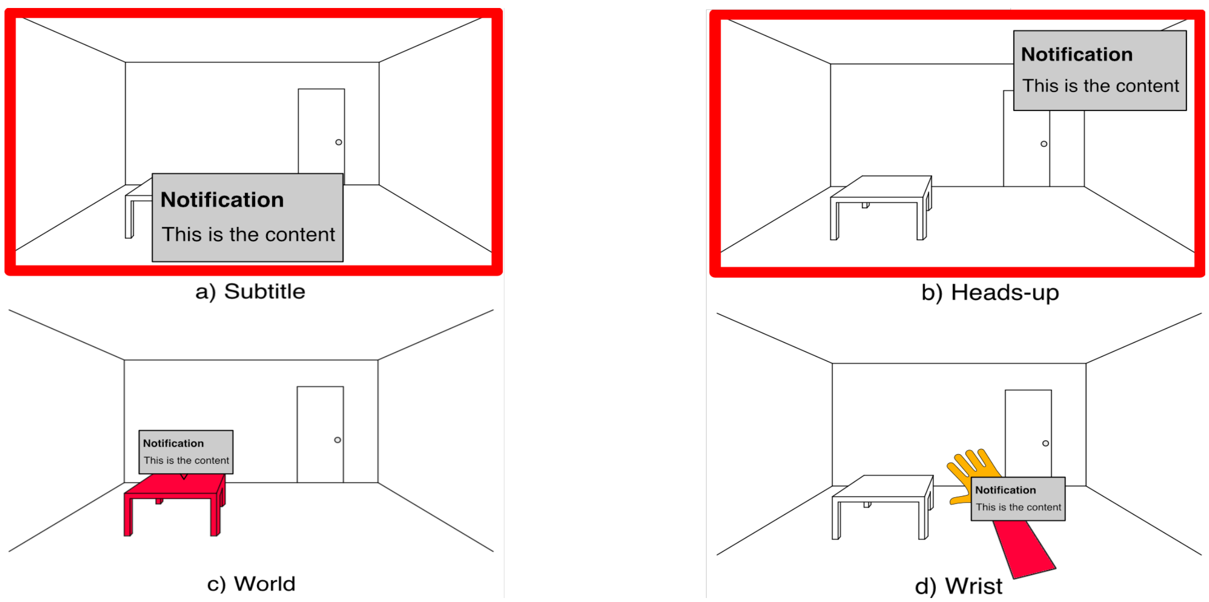


Figure 1: Different notification placements (location is illustrated by the red color, a) and b) are presented in screen space) : Subtitle, Heads-up. World and Wrist

ABSTRACT

Visual notifications are omnipresent in applications ranging from smart phones to Virtual Reality (VR) and Augmented Reality (AR) systems. They are especially useful in applications where users performing a primary task have to be interrupted to react to external events. However, these notifications can cause disruptive effects on the performance of users concerning their currently executed primary task. Also, different notification placements have been shown to have an influence on response times, as well as e.g. on user perceived intrusiveness and disruptiveness.

We investigated the effects and impacts of four visual notification types in AR environments when the main task was performed (1) in AR and (2) the real world. We used subtitle, heads-up, world space, and user wrist as notification types. In a user study, we interrupted the execution of the main task with one of the AR notification types. When noticing a notification, users responded to it by completing a secondary task. We used a Memory card game as the main task and the pressing of a correctly colored button as the secondary task. Our findings suggest that notifications at a user's *wrist* are most suitable when other AR elements are present. Notifications displayed in the *World* are quick to notice and understand if the view direction of a user is known. *Heads-up* notifications in the corner of the field-of-view, as they are primarily used in smart glasses, performed significantly worse, especially compared to *Subtitle* placement. Hence, we recommend to use different notification types depending on the overall structure of an AR system.

SUI '22, December 1–2, 2022, Online, CA, USA

ACM ISBN 978-1-4503-9948-7/22/12…$15.00
https://doi.org/10.1145/3565970.3567701

CCS CONCEPTS

• **Human-centered computing → User studies**; *Mixed / augmented reality*.

KEYWORDS

augmented reality, notifications, attention

ACM Reference Format:
Lucas Plabst, Sebastian Oberdörfer, Francisco Ortega, and Florian Niebling. 2022. Push the Red Button: Comparing Notification Placement with Augmented and Non-Augmented Tasks in AR. In *Symposium on Spatial User Interaction (SUI '22), December 1–2, 2022, Online, CA, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3565970.3567701

1 INTRODUCTION

Visual notifications in AR can be employed to draw a user's attention away from their current main task, towards specific, potentially important events. This is especially significant in safety-critical environments such as control rooms, medical care, disaster response, or construction, where missed incidents can sometimes have fatal consequences. Nevertheless, interruptions have disruptive effects on the user's task performance and lead to a higher memory load of users at the time of interruption [3]. Also, different forms of presentations in mobile device applications have an influence on response time and the disruption perceived from a notification [27, 35]. Rzayev et al. show that placement has an impact on the perceived urgency and intrusiveness of visual notifications in AR [32]. In Virtual Reality, presentation and placement of notifications have also been shown to influence response time, noticeability, distraction, intrusiveness [33], as well as perceived disruptiveness [17]. Thus far, the effects of notifications and notification placement on a main task are well established for desktop, mobile applications and VR. In contrast, research mostly focused on the effects of AR-based notifications on the performance of real world tasks. This, however, leaves out situations in which AR also displays the main task in addition to the notifications. Here, the perception of notifications can be vastly different depending on the amount of virtual content displayed. An AR-based notification during a real world task can stand out more than during the purely virtual task, e.g., due to the vergence-accommodation conflict, latency, differing color appearance and depth perception, especially when using Optical See-Through devices.

In this work, we focus on the perception and notability of four different AR-based notifications displayed either during a real world only or an AR-based task: subtitle, heads-up, world space, and user wrist. In particular, we investigate which notification type is more suited when (1) the only virtual information are notifications and (2) the virtual information is used to display the task and the notifications. The main task consists of a card game known as *Concentration* or *Memory* [37], where users have to find matching pairs of cards that are initially laid out face down on a surface. In the real world condition, no virtual content besides the visual notifications is presented to the user. In the AR condition, the card game itself is performed in AR. In both conditions, users have to interrupt their main task to perform an additional activity, i.e., pressing the button mentioned in the notification. We report on task performance between the conditions concerning *e.g.* reaction time, missed notifications, and error rate in the interrupting activity.

Our main contributions are foundations for AR notification display location, including (1) understanding the ideal position for AR notifications, which includes the effect the position of a notification in AR has on the main task performance and the effect the position has on the perception of the notification; (2) how these effects change depending on whether the main task is of a physical or fully virtual nature.

2 RELATED WORK

2.1 Notifications

A notification has been defined as a visual cue, auditory signal, or haptic alert generated by an application or service that relays information to a user outside of the current focus of attention [19]. Notifications have become an essential part of our interaction with technology on a day-to-day basis, especially with the unprecedented rise of the smartphone [7]. While difficult to accurately pin down, some studies have found their participants to receive around 80 notifications per day, with some receiving up to 200 [1]. Given the amount, it is reasonable to assume that not all notifications always arrive at an appropriate time for the user, which is problematic as Stothart et al. [36] has shown, that receiving a notification can significantly decrease the performance of an attention-demanding task. Several researchers have tried to manage the attentional cost of notifications by approaches such as grouping many notifications together in small batches delivered multiple times throughout the day [9] or by developing context aware delivery systems [30]. However, not receiving notifications can lead to increased frustration and actually lower productivity [22]. Also, not every notification can even be delayed until a later time, for example phone call notifications or time-critical alerts like in safety-critical-systems need to be delivered regardless of opportune timing or context. Orlosky et al. [29] have shown that the use of a head-mounted display for notification delivery can lead to increased spatial awareness with minimal performance impact over the use of a smartphone.

2.2 Information Acquisition in 3D

When it comes to placing content in AR environments, there are three possibilities according to the classification of Billinghurst et al.[4]:

- Head-stabilized: Information is fixed to the user's viewpoint.
- Body-stabilized: Information is fixed to the user's body.
- World-stabilized: Information is fixed to real-world locations.

Rzayev et. al. [33] experimented with different notification positions in VR. They concluded that there was not a preferred notification placement for all contexts, as each position was perceived differently from the others, but rather, that position should depend on the context of the notification and the current task the user is performing. Also researching notifications in VR, Ghosh et al. [12] explored interruptions and notifications in VR with several modalities like haptics and audio and derived design guidelines based on their findings. To evaluate the perception of notifications they created several questions, which will also be used during the course of this work. Lu et al. [25] developed an interface for quickly accessing short information at the periphery of vision using different glancing methods, which could be employed for notifications as well. Chua et al. [5] investigated the display-position of a monocular head-mounted-display and how it affected the performance and usability in a dual-task scenario. They found that middle-right, top-center, and top-right are most suited when the center of vision is needed for the main task and when the secondary stimulus is not urgent. Middle-center and bottom-center positions were preferred when the secondary stimulus required high noticeability. Rzayev et al. [32] also looked at notification position in AR during social interactions and found that displaying notifications in the user's field-of-view (FOV) was seen as favorable in social interactions. Participants could not agree on whether they preferred a center or top right position. Also based on a more casual day-to-day activity like talking to another person, Lucero et al. [26] developed and researched notifications on an AR-headset while walking and performing a pedestrian navigation task in a busy city center. They used a minimal UI and a discrete thumb touch-pad device to control notifications and found that participants had little issue with dealing with the notifications while being exposed to potential hazards in an urban environment. This might change with the increase in AR-content displayed, as it was shown that more virtual objects in an AR scene decreased task performance due to clutter [10].

2.3 Text in AR

Unlike traditional displays, the background of digital content on AR headsets can not be completely freely chosen, as it largely depends on the background of the surrounding. Especially text legibility decreases or increases greatly depending on the contrast ratio [24] and with AR headsets, the contrast ratio cannot always be kept constant. Debernadis et. al. [6] evaluated the presentation of text on different AR-headsets. They found that presenting text on a dark blue billboard with white text seems to be a good combination for indoor AR-applications, regardless of device or background. If the notification should also convey information through color, using the color as a background with white text is preferable. This was also corroborated by Gabbard et al. [11], who found a billboard-style with a semi-transparent background to be the most well-suited text display style across several outdoor background textures. Rzayev et al. [34] presented a study to evaluate reading text on an AR headset, looking especially at the positioning of text and presentation method (Rapid Serial Visual Presentation (RSVP) and line-by-line scrolling). The positions researched were center, upper right, and bottom-center. When text was displayed in the lower-center or center position, comprehension increased while perceived workload decreased, with participants preferring bottom-center for reading. Text in the top right was least favorable but might be suited for quick alarms or notifications, as longer reading caused eyestrain and reduced text comprehension.

3 EXPERIMENT

We conducted an experiment to determine if the position of a notification could affect primary and secondary task performance and to also examine if the perception of the notification changes depending on its location. In this study, participants were instructed to play a card game (primary task), during which they received notifications on an optical see-through AR headset, the Microsoft Hololens 2, to which they had to respond to (secondary task). The headset features a resolution of 1440x936 pixels per eye with a FOV of 43 degree horizontal, 29 vertical and 52 diagonal.

3.1 Design

In the experiment, we deployed four different notification PLACEMENTS (see Figure 1) in an AR-environment with two different TASK scenarios. This resulted in a mixed group design with two independent variables. The notification PLACEMENTS consisted of a notification in the (1) top right *(heads-up-display)*, (2) bottom middle *(subtitle)* portion of the AR headset display, (3) projected on the *wrist*, and (4) situated above the TASK in the *world*. The participants were exposed to all notification types during the experiment, i.e., type of notification was a within-subjects condition.

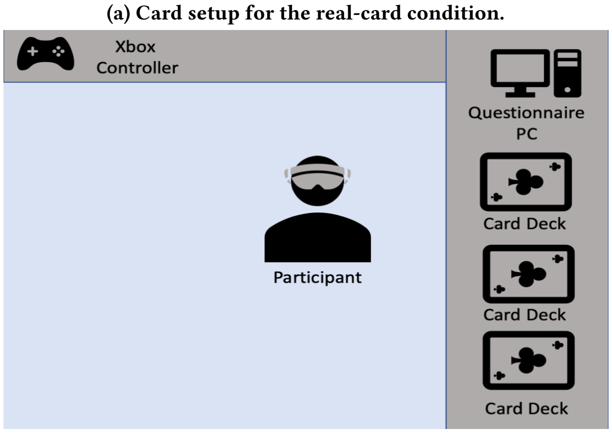
Cards were either all physical (REAL CARD TASK) or exclusively virtual AR playing cards (AR CARD TASK), depending on the TASK scenario. This between-subject variable allowed for a comparison of whether more virtual content influences notification perception. Also, this approach enabled an investigation on the influence of the FOV of the AR device. The real card condition was not limited to the boundaries of the AR device's FOV and resembled a typical task that could be enhanced with AR information like assisting in safety-critical medical procedures [31]. In contrast, the digital card condition explored the effectiveness of notifications in a more AR-focused situation.

The TASK for the experiment was a memory card-playing game, as a sustained attention task was needed for the experiment. Because this game requires a lot of recall ability, intrusive interruptions should have a large impact on the performance, which makes this task suited for investigating the effects of the notifications. The game rules are also very simple so preexisting knowledge about the game should not be an issue.

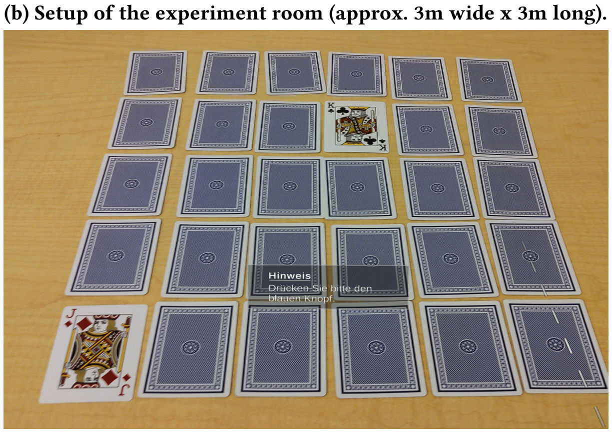
Participants were given three playing-card decks (cut down to 15 pairs from 28) spread out face-down in a five by six grid each (see Figure 2a). The decks were kept small to allow displaying all cards within the FOV of the AR headset used. Both card types measured 64mm X 89mm. Two cards had to be flipped and discarded if their color (red or black) and value matched. If they did not match, they had to be returned face-down. This process was repeated until all cards of a deck had been discarded, and then the participant had to move on to the next deck. Users playing with the digital cards could use their right index finger to tap on a card, which would flip it. If the cards matched, the cards would automatically disappear, or in the case of a mismatch, be flipped face-down again.

While playing the memory game, notifications with different PLACEMENTS were shown to the participants in the AR-headset. Each experiment run lasted eight minutes and the notifications appeared every 50 seconds resulting in a total of nine notifications per run. Timings were kept constant in an attempt to reduce potential confounds. Timing could influence the perception and we want to research this in the future. Ending each experiment run after eight minutes ensured every participant got shown exactly the same amount of notifications. The amount of total cards was chosen to make sure no participant could finish within the given

(a) Card setup for the real-card condition.



(b) Setup of the experiment room (approx. 3m wide x 3m long).



(c) Real Cards with notification. (Image has lower contrast due to capturing technique used by Hololens.

Figure 2: Experiment setup

time, which would cause them to sit idle, most likely influencing the results.

Every notification contained an instruction, which the participant needed to perform. Each instruction was the pushing of a specific button on an Xbox game-controller located on the far end of the room. The controller was placed away from the participant as it would cause a more severe distraction from the main task because it required the participants to completely stop the main task in order to attend to the instruction, resembling for example the dismissal of a patient monitoring system alarm in emergency health care. The four face-buttons on the controller are colored (yellow, red, blue, green), so each instruction told the participant to press a specific colored button, which was chosen at random. Notifications were not able to be dismissed by the user but disappeared after five seconds on their own, no matter if a correct button was pressed. The Xbox controller was paired with the Hololens 2 using a Bluetooth connection and plugged into a USB charger to avoid power loss during the experiment. Because of the power cable, and also to discourage participants from taking the controller to their seat, the controller was taped down. The study had a duration of approximately 65 minutes per participant.

3.2 Notification Design

All notifications had a rectangular form, mimicking the alerts most commonly seen on mobile and desktop operating systems. They featured a bold title, which read "Notice" for all notifications and a text-body, containing the instruction to be carried out. To ensure high legibility of the text, a dark gray background was chosen, along with white text color, keeping in line with design recommendations by Microsoft [28] and Jankowski et al. [20]. The font size was set to 20pt, putting it within range of Microsoft's guidelines for text legibility in AR applications. All notifications automatically aligned to face the user, with the exception of the z-axis(roll), therefore ignoring head tilting. It was found that as focal switching distance increased in AR, eye fatigue increased while performance decreased [2], so the notifications are displayed at about the same distance as the cards.

3.2.1 Heads-up. The notification types *Heads-up* and *Subtitle* are fixed to a specific position in the display of the headset and are therefore head-stabilized according to the categorization by Billinghurst and Kato [4]. As such they are visible regardless of position and orientation of the user and resemble traditional 3D user interfaces found in for example video games. Aside from the location in the display, both notifications are identical. The heads-up notification (see Figure 3a) is placed at the top-right border of the FOV, closely mimicking headsets such as Google Glass. Participants in Rzayev et al.'s [32] study disliked this position for longer reading but expressed that it would be well suited for short texts.

3.2.2 Subtitle. The subtitle notification (see Figure 3b) is placed at the bottom center border of the FOV, as suggested by Chua et al. [5] for dual-task scenarios that require high noticeability on the secondary stimuli. The bottom-center was chosen rather than middle-center, because we wanted to have the least visibility-impact on the main task, while also having good noticeability of the notification. Work by Rzayev et al. [34] also showed that participants preferred the bottom-center over the middle-center notification position when it came to reading text in AR. System notifications that are displayed by the Windows OS running on the Hololens 2 are displayed using the subtitle notification placement. With content placed close to the user, special attention needs to be paid to the vergence-accommodation-conflict [16]. The Hololens 2 display is fixed at an optical distance of approximately two meters away from the user, so Microsoft recommends placing content close to this point for extended interactions, with the optimal zone being one to five meters. Both *Subtitle* and *Heads-up* notifications are displayed at a distance of one meter away from the user, in accordance with the comfort guidelines by Microsoft. The notifications do not move in depth and the Hololens 2 automatically calibrates the interpupillary distance (starting with OS Version 20H2 released Nov. 2020) , which both lessens the potential discomfort caused by the vergence-accommodation-conflict.

3.2.3 Wrist. The body-stabilized notifications (see Figure 3c) were positioned at the user's right wrist, inspired by notifications that a user wearing a smartwatch might receive. Unlike a smartwatch, rotating the wrist did not change the position of the notification as it always remained centered above the wrist, independent of rotation.

3.2.4 World. The world-stabilized notifications (see Figure 3d) were placed on the center top edge of the card deck, which the user is currently closest to, hovering slightly over the table to not cause any alignment issues with the real world or block sight of the cards. It was found that world-stabilized interfaces centrally and closely located to the task improved task completion time [18].

3.3 Implementation

To run the experiment, a prototype was developed which could periodically send different notifications and allowed the user to play a memory-card-game. The application was implemented using Unity Engine 2020.3.16f1 with the help of Microsoft's Mixed Reality Toolkit (MRTK) v2.4.0. Aligning the digital playing cards with the table, as well as anchoring the world-notifications above the playing cards, was realized using Vuforia Image target recognition. Three tracking markers were generated using the ARMaker Tool [23] and printed on white paper with a size of 21cm x 21cm. Each deck had its own marker for tracking, as this reduced any shift that might have occurred by small registration errors. Every event (notification sent, button press, card flips and card matches) was logged by the application running on the Hololens 2 and saved to a text-file for analysis. Digital cards could be turned over by touching them with a small orange orb attached to the tip of the index finger on the right hand for additional visual feedback.

3.4 Measurements

We used the following measures to evaluate the notifications.

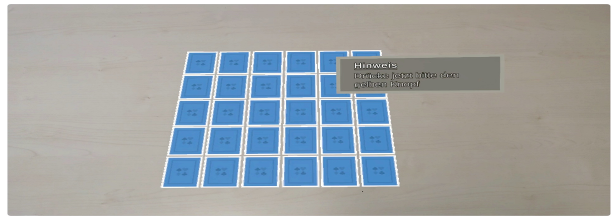
Performance. When a notification was sent and the participant subsequently reacted with pressing the corresponding button on the controller, it was measured as a *correct button press.* If a notification was sent and no button was pressed, the notification was counted as missed after 20 seconds of popping up. *Missed notification count* was measured as a dependent variable. In the event the button pressed did not match the instruction, it was counted as a *wrong button press.* Other measurements done during the experiment were the time after a notification was sent until a button was pressed (*reaction time*) and the number of *correct card matches.* As each experiment run is capped at eight minutes, the correct card matches are used as a measure of main task performance, instead of time to completion.

Usability and Task Load. After each round of the TASK, participants had to complete a set of questionnaires. To assess the overall usability of the notifications, the System Usability Scale (SUS) [13] was used, along with a NASA Task-Load-Index questionnaire (NASA-TLX) [15] to assess task load. Subscales of the NASA-TLX questionnaire were not weighted (Raw-TLX) as doing so does not seem to impact the results [14].

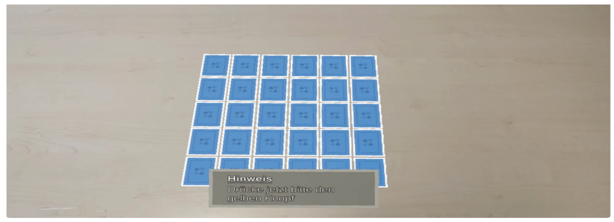
Perception. To gain a further understanding of the perception of notifications we used the questions from the work of Ghosh et al. [12] about noticeability, understandability, urgency and intrusiveness of notifications. In addition to the questionnaires, the participants were asked to rank each notification type by personal preference and give a short explanation of their ranking.

3.5 Procedure

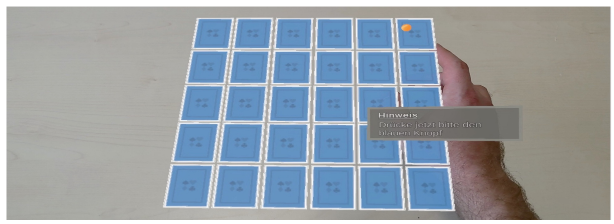
Task and order of notification Placements were assigned to each participant at random, so everyone would play with either real or digital cards with every notification type. Also, lighting conditions



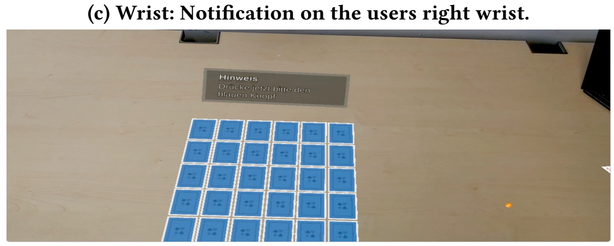
(a) Heads-up: Notification in top-right.



(b) Subtitle: Notification in bottom-center.



(c) Wrist: Notification on the users right wrist.



(d) World: Notification over the current task.

Figure 3: Different notification placements used in the experiment, here with the AR-Task.

were kept the same across all conditions and participants. After welcoming the participants, we asked them to read and sign experimental consent forms and fill out demographic information questionnaires. Following that, we explained the study procedure and explained how to correctly put on the Hololens 2. Participants had the rules of the game explained to them and were instructed to read the notifications and act upon the instructions contained in them. They were then given the headset and when the participant correctly put it on, we played a demo scene to show what kind of notification types the participant could expect and to ensure that each notification type was legible to the participant. The Hololens 2 automatically calibrates the display to the wearer's interpupillary distance after wearing it for about 30 seconds, so no separate calibration was done, as the process would automatically complete during the demo, ensuring optimal clarity for every participant. As soon as the demo finished, a blue text box was displayed, telling the participants that the experiment would start as soon as the text box disappeared. Also, over each tracking marker a purple cube was displayed to indicate if the marker was being correctly tracked. To set the different experiment conditions before each experiment run, the buttons on the Xbox Controller were used. When the participant confirmed that all markers were being tracked, the experiment was started by experiment leader by also pressing a button on the controller. The participant would then start playing the card game for eight minutes, with a pop-up text box notifying the participant of the end of the experiment when that time had passed. After each game had finished, the participant was instructed to take off the headset and complete a set of questionnaires. This procedure was then repeated (with the omission of the demo), so each participant would perform the Task a total of four times, once for every notification type. Participants had to sit down during the experiment and only got up from their chair to press the button on the controller.

The institutional review board of Human-Computer-Media Würzburg approved our ethics proposal for this study.

3.6 Participants

Participants were recruited from a pool of university students studying Human-Computer-Systems or Media-Communication. They are required to gather a certain amount of experiment participation hours for their coursework and were rewarded with 1.25 hours participation time for the experiment. In total, 40 participants were recruited (12 Male and 28 female). Age ranged from 19 to 30 years (M = 22.15, SD = 2.3). All either had normal or corrected-to-normal vision. Of those participants, all 40 stated that they used smartphones and the internet daily and 30 had either never used AR before or only in experiments (16 in experiments, 14 never). Additionally, 11 participants stated that they played video games somewhat regularly and 36 were right-handed. The data of one participant could not be evaluated at all. Two participants were missing the log files from the playing card game, but their questionnaires were still evaluated. This brings the total to 39 participants for the questionnaires and 37 for the task. All participants were fully vaccinated against COVID-19 and were required to show a negative test that was taken at most 24 hours before their participation. During the experiment participants wore medical gloves which were disposed of after the experiment.

4 RESULTS

We analyzed our data with RStudio in version 1.4.1106. To compare the means of the conditions for the measured factors, because our data did not meet the assumptions for ANOVA, we transformed it using the Aligned Rank Transform[38], before computing a non-parametric ANOVA on the transformed data. Pairwise tests were done using the ART-C Procedure [8].

Most notable, we found that *Wrist* notifications performed substantially different depending on *Task*, that *Heads-up*- performed worse overall than *Subtitle*- notifications and that *World* notifications performed the best overall in most measurements. *World* notifications were ranked as the most preferable, while *Heads-up* notifications were ranked the lowest. In the following we will go through the results of the experiment in detail, starting with the quantitative performance measurements like reaction time, then moving on to results gathered from the questionnaires and qualitative results. Table 1 provides a summary of the descriptive statistics.

4.1 Performance

Correctly Pressed Buttons. We found a significant two-way interaction between Placement and Task explaining the number of correctly pressed buttons, $F(3, 105) = 5.57$, $p = 0.001$. We found a significant main effect for the Placement $F(3, 105) = 3.88$, $p = 0.011$. However we found no significant effect for the Task overall $F(1, 35) = 1.7849$,$p = 0.190$. Contrast tests showed a significant difference between *Subtitle* and *World* $t(105) = 3.12, p = 0.012$, and *Subtitle* and *Wrist* $t(105) = 2.62, p = 0.049$, regardless of Task. Given the AR card condition, we found significant differences between *World* (6.47) and *Wrist* (8.37), $t(105) = 3.71, p = 0.008$ with respect to the number of correctly pressed buttons, (see Figure 4a). We also found a significant difference for *Wrist* depending on Task $t(129.8) = 3.44, p = 0.017$. When conducting difference of difference testing, we found that the difference between *Subtitle* and *World* $t(105) = 3.12, p = 0.012$, as well as *Subtitle* and *Wrist* $t(105) = 2.61, p = 0.049$, changed significantly depending on the Task.

Wrongly Pressed Buttons. We found no significant effect between Placement and Task explaining the number of wrongly pressed buttons, $F(3, 105) = 0.311$, $p = 0.817$. There was also no main effect regarding Placement or Task.

Missed Notifications. We found a significant two-way interaction between Placement and Task explaining the number of completely missed notifications, $F(3, 105) = 5.09$, $p = 0.002$. We also found a significant main effect for the Placement , $F(3, 105) = 4.82$, $p = 0.003$.

Contrast tests showed a significant difference between *Subtitle*(0.97) and *World*(2.13), $t(105) = 3.57, p = 0.003$, and *Subtitle*(0.97) and *Wrist*(2.05) $t(105) = 2.91, p = 0.022$, regardless of Task. Of note, though not statistically significant, the remaining *Subtitle* comparison to *Wrist* had p = 0.0751. These results show that there may be a significant difference between subtitle and other type of tested notification regardless of Task. In the real card condition, we found significant differences between *Subtitle*(0.94) and *Wrist*(3.61) $t(105) = 3.42, p = 0.019$, with respect to the number of missed notifications regardless of cards. In the AR card condition, we

Table 1: Descriptive statistics; $N = 37$**, real cards** $n = 18$**, AR cards** $n = 19$**. Values are** $M(SD)$**.**

Scale	Overall	Heads-up	Subtitle	World	Wrist
Correctly Pressed Buttons (0-9)					
Real Cards	6.82 (2.8)	6.89 (2.54)	7.89 (2.27)	7.11 (2.03)	5.39 (3.70)
AR Cards	7.34 (2.36)	6.68 (2.79)	7.84 (2.01)	6.47 (2.44)	8.37 (1.67)
Wrongly Pressed Buttons (0-9)					
Real Cards	0.111 (0.358)	0.167 (0.383)	0.167 (0.514)	0.111 (0.323)	0.0 (0.0)
AR Cards	0.132 (0.411)	0.263 (0.562)	0.158 (0.501)	0.053 (0.229)	0.053 (0.229)
Missed Notifications (0-9)					
Real Cards	2.07 (2.84)	1.94 (2.58)	0.944 (2.29)	1.78 (2.02)	3.61 (3.70)
AR Cards	1.53 (2.28)	2.05 (2.72)	1.0 (1.92)	2.47 (2.29)	0.579 (1.68)
Reaction Time (in seconds)					
Real Cards	6.83 (1.58)	6.54 (0.92)	6.13 (0.84)	5.95 (0.56)	8.69 (1.88)
AR Cards	7 (1.68)	7.3 (1.72)	6.89 (1.65)	6.01 (1.08)	7.8 (1.75)
Correct Card Matches (0-45)					
Real Cards	23.4 (6.97)	23.4 (8.07)	24.7 (7.24)	24.3 (5.89)	21.3 (6.52)
AR Cards	24.6 (6.78)	21.4 (6.51)	24.7 (7.88)	28.5 (4.95)	23.8 (5.93)
SUS (0 - 100)					
Real Cards	71.1 (14.4)	67.2 (16.0)	75.1 (13.0)	71.7 (14.2)	70.3 (14.3)
AR Cards	67.9 (20.4)	62.5 (23.9)	70.6 (18.2)	70.1 (18.4)	68.4 (21.2)
NASA TLX (0-100)					
Real Cards	35.4 (15.1)	34.7 (15.5)	34.3 (14.4)	33.6 (14.7)	38.9 (16.4)
AR Cards	42.1 (17.7)	46.9 (19.7)	40.0 (18.7)	38.1 (16.8)	43.2 (15.2)
Noticeability (1-7)					
Real Cards	4.87 (1.84)	4.32 (1.89)	5.68 (1.42)	5.26 (1.79)	4.21 (1.9)
AR Cards	4.75 (1.90)	4.45 (2.21)	5.45 (1.73)	4.35 (1.93)	4.75 (1.62)
Understandability (1-7)					
Real Cards	5.71 (1.84)	5.21 (2.07)	6.42 (0.97)	6.26 (1.05)	4.95 (2.46)
AR Cards	5.74 (1.69)	4.9 (2.1)	5.8 (1.8)	6.2 (0.95)	6.05 (1.5)
Urgency (1-7)					
Real Cards	4.63 (1.30)	4.84 (1.12)	4.74 (1.37)	4.74 (1.2)	4.21 (1.51)
AR Cards	4.2 (1.72)	4.15 (1.93)	4.3 (1.62)	4.05 (1.76)	4.3 (1.69)
Intrusiveness (1-7)					
Real Cards	4.46 (1.12)	4.32 (0.95)	4.79 (0.92)	4.53 (0.96)	4.21 (1.55)
AR Cards	4.28 (1.37)	4.15 (1.5)	4.65 (0.99)	3.7 (1.49)	4.6 (1.31)

found significant differences between *World*(2.47) and *Wrist*(0.579) $t(105) = 3.91, p = 0.004$ and between *Subtitle*(1) and *World*(2.47) $t(105) = 3.11, p = 0.04$, with respect to the number of missed notifications. When conducting difference of difference testing, we found that the difference between *HUD* and *Wrist* $t(84) = 2.29, p = 0.024$, changed significantly depending on the TASK.

Reaction Time. Two participants didn't respond to any notifications so these participants are excluded from this calculation for reaction time. We only used notifications that produced a response to calculate reactions. We found a significant main effect for PLACEMENT, $F(3, 84) = 27.66$, $p < 0.001$. However we found no significant effect for the TASK $F(1, 28) = 0.16$,$p = 0.069$ or for an interaction between the PLACEMENT and TASK explaining the reaction time $F(3, 84) = 2.27$, $p = 0.08$. Contrast tests revealed significant differences between every PLACEMENT $p < 0.001$ except *Heads-up* and *Subtitle* $p = 0.17$, disregarding the TASK. Given the real card condition, we found significant differences between *Heads-up*(6.54) and *Wrist*(8.69) $p < 0.001$, between *Subtitle*(6.13) and *Wrist*(8.69), $p < 0.001$, and between *World*(5.95) and *Wrist*(8.69) $p < 0.001$ with respect to the reaction time (see Figure 4c). Given the AR card condition, we found significant differences between *Heads-up*(7.3) and *World*(6.01) $p < 0.001$ and between *Wrist*(7.8) and *World*(6.01) $p < 0.001$ with respect to the reaction time (see Figure 4c). When conducting difference of difference testing, we found a significant difference between *Heads-up* and *Wrist* $t(105) = 2.29, p = 0.024$,*Subtitle* and *Wrist* $t(105) = 3.03, p = 0.03$ and *World* and *Wrist* $t(105) = 3.47, p < 0.001$ depending on the TASK.

Correct Card Matches. We found a significant two-way interaction between Placement and Task explaining the number of correctly matched cards, $F(3, 108) = 2.75$, $p = 0.046$. We found a significant main effect for Placement, $F(3, 108) = 5.87$, $p < 0.001$. However, the Task caused no significant main effect, $F(1, 36) = 0.71$, $p = 0.40$. Pairwise contrast tests revealed significant differences between *Heads-up*(22.4) and *World*(26.4) , $t(108) = 3.36, p = 0.006$ and between *Wrist*(22.55) and *World*(26.4) , $t(108) = 3.64, p = 0.002$, with respect to the correct card matches regardless of Task. In the AR condition we found a significant difference between *Heads-up* and *World* $t(108) = 4.18, p = 0.001$. When conducting difference of difference testing, we found that the difference between *Subtitle* and *World* $t(108) = 2.72, p = 0.007$, changed significantly depending on the Task.

4.2 Questionnaires

Usability. We found a significant main effect for Placement in explaining the Usability score, $F(3, 111) = 3.43$, $p = 0.019$. However, the Task caused no significant main effect, $F(1, 37) = 0.09$, $p = 0.76$. Pairwise contrast tests revealed significant differences between *Heads-up*(64.85) and *Subtitle*(72.85) with respect to the SUS score regardless of Task, $p = 0.013$.

Task Load. We found a significant main effect for Placement in explaining the task load, $F(3, 111) = 4.29$, $p = 0.006$. However, the Task caused no main effect, $F(1, 37) = 0.93$, $p = 0.33$. Pairwise contrast tests revealed significant differences between *Heads-up*(40.8) and *World*(35.85) , $t(111) = 2.63, p = 0.047$, and *Wrist*(41.05) and *World*(35.85) , $t(111) = 3.04, p = 0.015$, with respect to the task load regardless of Task.

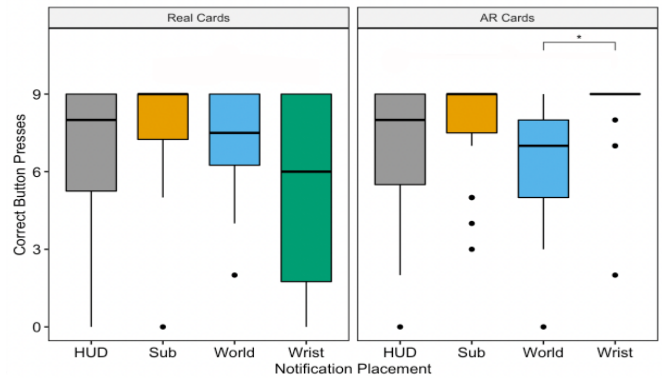
4.2.1 Perception.

Noticeability. We found a significant main effect for the Placement in assessing the noticeability, $F(3, 111) = 3.58$, $p = 0.016$. However, the Task caused no main effect, $F(1, 37) = 0.27$, $p = 0.604$. Pairwise contrast tests revealed significant differences between *Heads-up*(4.38) and *Subtitle*(5.56) , $t(111) = 2.79, p = 0.03$, and *Wrist*(4.48) and *Subtitle*(5.56), $t(111) = 2.88.p = 0.024$, with respect to the perceived noticeability regardless of Task.

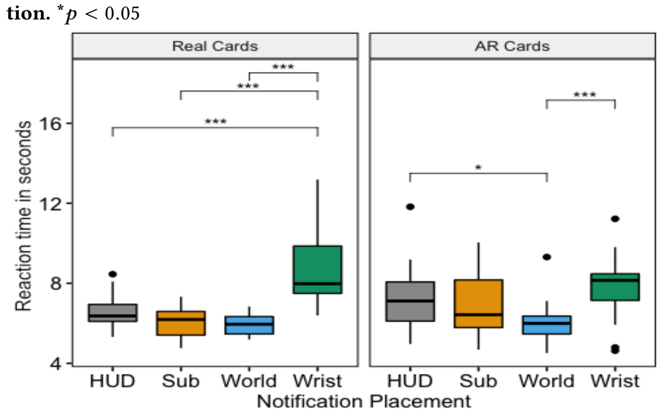
Understandability. We found a significant main effect for Placement in assessing the understandability, $F(3, 111) = 3.67$, $p = 0.014$. However, the Task caused no main effect, $F(1, 37) = 0.46$, $p = 0.49$. Pairwise contrast tests revealed significant differences between *Heads-up*(5.05) and *Subtitle*(6.11) $t(111) = 2.74, p = 0.035$, and *Heads-up*(5.05) and *World*(6.23) , $t(111) = 2.88, p = 0.025$, with respect to the perceived understandability regardless of Task.

Urgency. We did not find any significant effects for the Task $F(1, 37) = 0.147$, $p = 0.70$, Placement $F(3, 111) = 0.317$, $p = 0.812$ or the interaction between the two $F(3, 111) = 0.591$, $p = 0.621$ regarding the perceived urgency.

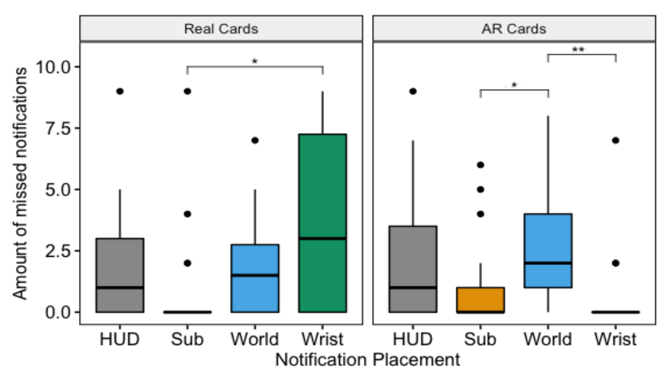
Intrusiveness. We found a significant main effect for the Placement in assessing the intrusiveness, $F(3, 111) = 2.77$, $p = 0.044$. We did not find any significant effects for the Task $F(1, 37) = 0.99$, $p = 0.32$ or the interaction between the two $F(3, 111) = 2.33$, $p = 0.078$ regarding the perceived intrusiveness. Pairwise contrast tests revealed significant differences between *World*(4.12) and *Subtitle*(4.72) $t(111) = 2.74, p = 0.035$ with respect to the perceived intrusiveness regardless of Task.



(a) Correct button presses: Depending on notification and card condition. $^{*}p < 0.05$

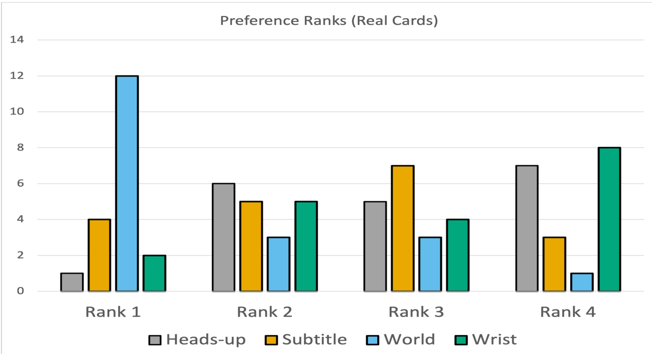


(b) Reaction time : Depending on notification and card condition. $^{*}p < 0.05$, $^{***}\ p < 0.001$

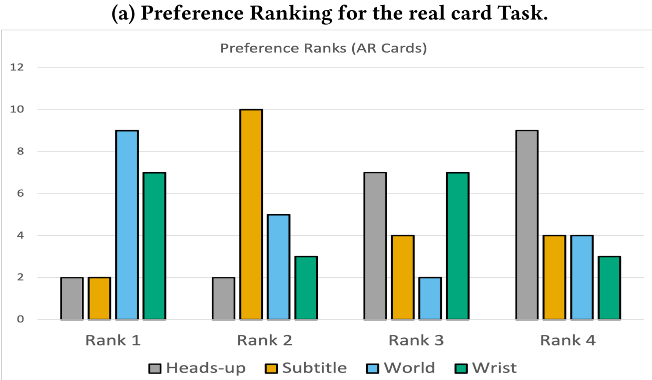


(c) Missed notifications: Depending on notification and card condition. $^{*}p < 0.05$,$^{**}\ p < 0.01$

Figure 4: Pairwise tests with Notification condition regarding correct button presses, reaction time and missed notifications.



(a) Preference Ranking for the real card Task.



(b) Preference Ranking for the AR-card Task.

Figure 5: Preference rankings for each notification type depending on the Task.

4.2.2 Subjective measures. At the end of the experiment, participants were asked to rank the Placement by preference. Ranks were weighted, so a Rank 1 preference would be worth 1 point and Rank 4 is 4 points. This number was then divided by the number of responses to reach an average ranking, with 1 being the most favorable ranking and 4 the least, which can be seen in Table 2 and Figure 5. Across both Task conditions, *World* notifications were ranked statistically significantly higher than the others ($p < 0.05$).

Table 2: Descriptive statistics; $N = 39$, real cards $n = 19$, AR cards $n = 20$. Values are average rank from 1 to 4 (1 being most favorable).

Scale	Real Cards	AR Cards	Total
Heads-up	2.95	3.32	3.05
Subtitle	2.47	2.63	2.49
World	1.63	2.16	1.85
Wrist	2.95	2.42	2.62

5 DISCUSSION

This study revealed significant effects of Placement and interaction effects between Placement and Task.

Wrist produced the best and also the worst score regarding the correctly pressed buttons, depending on Task. With the AR cards, *Wrist* produced more correct button presses than *Heads-up* or *World* and less missed notifications than *World*. The picture changes quite a bit when looking at the real card Task. Correct button matches, missed notifications and noticeability were all worse than *Subtitle* and reaction time was worse than all other Placements. When looking at the ranking responses, it can also be seen that the preference of *Wrist* notifications depends largely on the card condition, going from tied to last with real cards to second place with AR cards. This can be explained by the fact that with the real cards it was possible to perform the Task without looking through the FOV of the AR headset. Participants often glanced under the center of the display, which caused them to miss notifications on the *Wrist*, with two participants completely missing all *Wrist* notifications. When playing with AR cards, the user was forced to look through the display to see the cards, which ensured that they would also have their *Wrist* in the FOV when tapping on the cards. The other notification placements were largely unaffected by this behavior as they were rendered anyway. This is also supported by research done by Kruijff et al. [21], where they found lower FOV to negatively impact discovery rates of target objects, as they enter the FOV less often. Some participants with real cards explained their ranking with statements like "I did not pay attention to my wrist at all" which explains why they missed some of the notifications. *Wrist* notifications were also rated as the hardest to notice in the questionnaire.

World was better in reaction time in the AR Task than any other and better in the real cards as *Wrist*. They were also better in correct card matches and TLX score than *Heads-up* or *Wrist* across and showed better understandability than *Heads-up* across both Tasks. *World* was the most preferred position for notification Placement, independent of card. This was explained with statements like "over the playing field the notifications were the clearest and quickest to read". The decreased reaction time can be explained because having the notification positioned at their current focus point might lead to a quicker registration than when the user first needs to look somewhere else. Because the participants were quicker to respond, they also had more time to play the card game, leading to a higher match count. Beside the quicker reaction time in AR, *World* always performed the same as *Subtitle* with no significant differences between the two in any other measurement except rank, where *World* placed first overall, and *Subtitle* second.

Subtitle notifications had higher correct button presses and reaction time than *Wrist* with real cards and scored higher in the SUS than *Heads-up*. They were also evaluated as having a higher understandability as *Heads-up* and better noticeability than *Heads-up* and *Wrist*, regardless of Task. The results suggest that *Subtitle* notifications were relatively robust, providing consistent results in all measurements. They didn't require a large shift of attention away from the focus and had good legibility. Being attached to the user's FOV, it was impossible to move them out of vision.

Heads-up placement had a better reaction time and missed notification count than *Wrist* in the real card Task. Comparing *Heads-up* and *Subtitle* notifications, it can be seen that *Heads-up* did not perform better than *Subtitle* in any of the measurements, while also performing worse in several categories and placing last in the preference ranking. Both feature the advantage of always being in view, but participants stated that they found shifting attention to the top-right caused a higher distraction from the game. However, the *Subtitle* notifications enabled them to keep concentrating on the game, while simultaneously reading the notification. One participant stated "bottom-center was the easiest to view, because you're not completely distracted from the game, but you still notice that there is a notification." *Heads-up* notifications were also more prone to an incorrect fit of the headset. Because they were in the peripheral vision not only vertically but also horizontally, wearing the headset incorrectly caused them to be cut off much more easily.

5.1 Design Recommendations

Use Wrist notifications in scenarios with high amount of interaction with AR content. These notifications can provide great results when the user is interacting with other AR content using their hands, but can be easily missed when the hands are outside of the FOV. If the user is not interacting with any other AR content in the environment, other notifications are better suited. Example: Notifications while the user is modifying a 3D model.

Use World notifications if you know where the user is going to look or for stationary tasks. *World* notifications perform well when placed in close proximity to the current task the user is focusing on, without blocking view of the task. If a user is stationary and focusing on a single area for a longer period of time, world notifications seem to be the preferred type of notification. However, if the user is shifting their attention or moving a lot, they might miss stationary notifications placed in the *World*. For these cases *Subtitle* notifications should be used. Example: Notifications while the user is sitting at a desk working on a computer monitor.

When in doubt, use Subtitle notifications for general notifications. *Subtitle* notifications can be displayed to the user regardless of context, do not require any outside world tracking, and are hardest to miss, while still being comfortable and easy to read. Example: Notifications while the user is moving around.

5.2 Limitations

Rating the number of matches might not be a good indicator of task performance, as playing speed and tactic varied between participants. Some prefer to play very systematically and avoid incorrect matches in favor of taking longer to memorize the card order, while others opted for a more direct approach and simply tried to flip cards as fast as possible. Using the real playing cards caused some participants to glance under the headset's display to flip cards as they weren't forced to look through the display, which caused them to completely miss *Wrist* notifications. This is a limitation of the headset's FOV and not necessarily with the placement. However, it is not clear that a larger FOV would have changed the result. Even though the task we chose was a very near-view environment, we still noticed the impact of the small FOV. Testing this in a bigger environment might increase the impact of the smaller FOV even more. But this means that hardware has to be considered when deciding on a certain notification type. Participants were also instructed to use only their right hand in both Task scenarios, as flipping the cards was only possible using the right index finger in the digital card condition. *Wrist*-notifications also only spawned on the right wrist. Even though participants were told this, some used their left hand or both hands during the real card condition. Another possible limitation was the context of the notifications. Participants were instructed to look at and carry out each instruction, which means that there were no unimportant notifications and participants knew to pay attention. This is reflected in the non-significant differences in urgency. The experiment was also set in a quiet and well-lit environment, largely free of any distractions, which might not necessarily reflect real-world conditions. The notifications also arrived at a fixed rate of every 50 seconds, which might have caused the participants to expect their arrival, although data does not seem to support this. Future work should research notification timing and frequency.

5.3 Future Work

Assisting notification delivery in AR through audio cues or haptic feedback should be researched. We saw that *wrist* notifications were missed more because it was possible that the wrist was outside of the headset's FOV. Using audio could help in noticing if a notification is currently present and draw attention to it. It is also worth researching whether notification indicators in the FOV, telling the user that a notification is pending, could improve the notification experience, as a user could then choose when to look at the notification. Notification placement should also be researched with a non-stationary task, especially as a walking task might feature different focal distances which could affect fatigue and performance. Issues with small FOV might also be increased in a non-near-view or non-stationary task. Another topic of interest is to repeat this study with a video-see-through headset and compare differences in notification perception between virtual- and augmented reality, and also optical-and video-see-through augmented reality. In our study the *Heads-up* notifications did not score significantly worse in Intrusiveness, while it did in a similar study done in VR [33], which might indicate that there are noteable differences between AR and VR notfications. Comparing the two could lead to a general notification design guideline in immersive 3D environments.

6 CONCLUSION

In this experiment we compared four different notification placements (*heads-up*, *subtitle*, *world*, *wrist*) in AR while performing one of two card gaming tasks containing physical playing cards *real cards*), or virtual playing cards (*AR-cards*) and constructed design recommendations for notifications in AR. We found that using notifications located on the *Wrist* should take into account how much interactivity or other content is present in the AR environment. Also when using head-stabilized notification in the user's periphery, bottom-center position should be used over top-right placement. The highest number of correct reactions to a notification, was present with *Wrist* notifications but only with a high amount of other virtual content in the environment. The quickest response to notifications was found with *World* notifications.

ACKNOWLEDGMENTS

This project is funded by: DFG–425868361 (SPP2199), NSF-1948254, NSF-2037417, NSF-2016714, NSF-2106590, ONR-N00014-21-1-2580, ONR-00014-21-1-2949N.

REFERENCES

[1] Utku Acer, Afra Mashhadi, Claudio Forlivesi, and Fahim Kawsar. 2015. *Energy Efficient Scheduling for Mobile Push Notifications*. Vol. 2. https://doi.org/10.4108/eai.22-7-2015.2260067 journalAbbreviation: EAI Endorsed Transactions on Energy Web.

[2] Mohammed Safayet Arefin, Nate Phillips, Alexander Plopski, Joseph L. Gabbard, and J. Edward Swan. 2022. The Effect of Context Switching, Focal Switching Distance, Binocular and Monocular Viewing, and Transient Focal Blur on Human Performance in Optical See-Through Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 28, 5 (May 2022), 2014–2025. https://doi.org/10.1109/TVCG.2022.3150503 Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[3] Brian P Bailey, Joseph A Konstan, and John V Carlis. 2000. Measuring the effects of interruptions on task performance in the user interface. In *Smc 2000 conference proceedings. 2000 ieee international conference on systems, man and cybernetics.'cybernetics evolving to systems, humans, organizations, and their complex interactions'(cat. no. 0*, Vol. 2. IEEE, 757–762.

[4] Mark Billinghurst and Hirokazu Kato. 1999. *Collaborative Mixed Reality*. Springer Berlin Heidelberg, 261–284. https://doi.org/10.1007/978-3-642-87512-0_15

[5] Soon Hau Chua, Simon T. Perrault, Denys J. C. Matthies, and Shengdong Zhao. 2016. Positioning Glass: Investigating Display Positions of Monocular Optical See-Through Head-Mounted Display. In *Proceedings of the Fourth International Symposium on Chinese CHI - ChineseCHI2016*. ACM Press, 1–6. https://doi.org/10.1145/2948708.2948713

[6] Saverio Debernardis, Michele Fiorentino, Michele Gattullo, Giuseppe Monno, and Antonio Uva. 2014. Text Readability in Head-Worn Displays: Color and Style Optimization in Video vs. Optical See-Through Devices. *IEEE transactions on visualization and computer graphics* 20 (Jan. 2014), 125–39. https://doi.org/10.1109/TVCG.2013.86

[7] Michael DeGusta. 2012. Are Smart Phones Spreading Faster than Any Technology in Human History? https://www.technologyreview.com/2012/05/09/186160/are-smart-phones-spreading-faster-than-any-technology-in-human-history/

[8] 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*. ACM, Virtual Event USA, 754–768. https://doi.org/10.1145/3472749.3474784

[9] Nicholas Fitz, Kostadin Kushlev, Ranjan Jagannathan, Terrel Lewis, Devang Paliwal, and Dan Ariely. 2019. Batching smartphone notifications can improve well-being. *Computers in Human Behavior* 101 (Dec 2019), 84–94. https://doi.org/10.1016/j.chb.2019.07.016

[10] Jonathan Flittner*, John Luksas, and Joseph L. Gabbard. 2020. Predicting User Performance in Augmented Reality User Interfaces with Image Analysis Algorithms. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 64, 1 (Dec 2020), 2108–2112. https://doi.org/10.1177/1071181320641511

[11] Joseph L. Gabbard, J. Edward Swan, and Deborah Hix. 2006. 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 (Feb. 2006), 16–32. https://doi.org/10.1162/pres.2006.15.1.16

[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* 24, 4 (April 2018), 1447–1456. https://doi.org/10.1109/TVCG.2018.2793698 Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[13] Rebecca A. Grier, Aaron Bangor, Philip Kortum, and S. Camille Peres. 2013. The System Usability Scale: Beyond Standard Usability Testing. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 57, 1 (2013), 187–191. https://doi.org/10.1177/1541931213571042 arXiv:https://doi.org/10.1177/1541931213571042

[14] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (Oct 2006), 904–908. https://doi.org/10.1177/154193120605000909

[15] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[16] David M. Hoffman, Ahna R. Girshick, Kurt Akeley, and Martin S. Banks. 2008. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision* 8, 3 (Mar 2008), 33. https://doi.org/10.1167/8.3.33

[17] Ching-Yu Hsieh, Yi-Shyuan Chiang, Hung-Yu Chiu, and Yung-Ju Chang. 2020. Bridging the Virtual and Real Worlds: A Preliminary Study of Messaging Notifications in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–14. https://doi.org/10.1145/3313831.3376228

[18] Samat Imamov, Daniel Monzel, and Wallace S. Lages. 2020. 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)*. 851–858. https://doi.org/10.1109/VR46266.2020.00110 ISSN: 2642-5254.

[19] 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 - CSCW '10*. ACM Press, Savannah, Georgia, USA, 27. https://doi.org/10.1145/1718918.1718926

[20] Jacek Jankowski, Krystian Samp, Izabela Irzynska, Marek Jozwowicz, and Stefan Decker. 2010. Integrating Text with Video and 3D Graphics: The Effects of Text Drawing Styles on Text Readability. (2010), 10.

[21] Ernst Kruijff, Jason Orlosky, Naohiro Kishishita, Christina Trepkowski, and Kiyoshi Kiyokawa. 2019. The Influence of Label Design on Search Performance and Noticeability in Wide Field of View Augmented Reality Displays. *IEEE Transactions on Visualization and Computer Graphics* 25, 9 (Sept. 2019), 2821–2837. https://doi.org/10.1109/TVCG.2018.2854737 Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[22] 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*. ACM, 1011–1020. https://doi.org/10.1145/2858036.2858359

[23] Shawn Lehner. 2017. ARMaker - Augmented Reality Marker Generator. https://shawnlehner.github.io/ARMaker/

[24] Chin-Chiuan Lin. 2003. Effects of contrast ratio and text color on visual performance with TFT-LCD. *International Journal of Industrial Ergonomics* 31, 2 (Feb 2003), 65–72. https://doi.org/10.1016/S0169-8141(02)00175-0

[25] Feiyu Lu, Shakiba Davari, Lee Lisle, Yuan Li, and Doug A. Bowman. 2020. Glanceable AR: Evaluating Information Access Methods for Head-Worn Augmented Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Atlanta, GA, USA, 930–939. https://doi.org/10.1109/VR46266.2020.00113

[26] Andrés Lucero and Akos Vetek. 2014. NotifEye: using interactive glasses to deal with notifications while walking in public. In *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology*. ACM, Funchal Portugal, 1–10. https://doi.org/10.1145/2663806.2663824

[27] Abhinav Mehrotra, Veljko Pejovic, Jo Vermeulen, Robert Hendley, and Mirco Musolesi. 2016. My phone and me: understanding people's receptivity to mobile notifications. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 1021–1032.

[28] Microsoft. 2022. Typography - Mixed Reality. https://learn.microsoft.com/en-us/windows/mixed-reality/design/typography

[29] Jason Orlosky, Kiyoshi Kiyokawa, and Haruo Takemura. 2014. Managing Mobile Text in Head Mounted Displays: Studies on Visual Preference and Text Placement. 18, 2 (2014), 12.

[30] Chunjong Park, Junsung Lim, Juho Kim, Sung-Ju Lee, and Dongman Lee. 2017. Don't Bother Me. I'm Socializing! A Breakpoint-Based Smartphone Notification System. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. Association for Computing Machinery, 541–554. https://doi.org/10.1145/2998181.2998189

[31] Lucas Plabst, Sebastian Oberdörfer, Oliver Happel, and Florian Niebling. 2021. Visualisation methods for patient monitoring in anaesthetic procedures using augmented reality. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*. ACM, 1–3. https://doi.org/10.1145/3489849.3489908

[32] 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*. ACM, Tallinn Estonia, 1–11. https://doi.org/10.1145/3419249.3420095

[33] 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*. ACM, Barcelona Spain, 199–211. https://doi.org/10.1145/3311350.3347190

[34] Rufat Rzayev, Paweł W. Woźniak, Tilman Dingler, and Niels Henze. 2018. 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*. ACM, Montreal QC Canada, 1–9. https://doi.org/10.1145/3173574.3173619

[35] Cary Stothart, Ainsley Mitchum, and Courtney Yehnert. 2015. The attentional cost of receiving a cell phone notification. *Journal of experimental psychology: human perception and performance* 41, 4 (2015), 893.

[36] Cary Stothart, Ainsley Mitchum, and Courtney Yehnert. 2015. The attentional cost of receiving a cell phone notification. *Journal of Experimental Psychology: Human Perception and Performance* 41, 4 (Aug. 2015), 893–897. https://doi.org/10.1037/xhp0000100

[37] Daniel J. Velleman and Gregory S. Warrington. 2013. What to Expect in a Game of Memory. *The American Mathematical Monthly* 120, 9 (Nov 2013), 787–805. https://doi.org/10.4169/amer.math.monthly.120.09.787

[38] 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*. ACM, Vancouver BC Canada, 143–146. https://doi.org/10.1145/1978942.1978963