Bridging the Virtual and Real Worlds: A Preliminary Study of Messaging Notifications in Virtual Reality

Ching-Yu Hsieh[1], Yi-Shyuan Chiang[2], Hung-Yu Chiu[1], Yung-Ju Chang[1]
[1]National Chiao Tung University, Hsinchu, Taiwan
[2]National Tsing Hua University, Hsinchu, Taiwan
{chingyu4206, yschiangg, hgy2395872}@gmail.com, armuro@nctu.edu.tw

ABSTRACT

Virtual reality (VR) platforms provide their users with immersive virtual environments, but disconnect them from real-world events. The increasing length of VR sessions can therefore be expected to boost users' needs to obtain information about external occurrences such as message arrival. Yet, how and when to present these real-world notifications to users engaged in VR activities remains underexplored. We conducted an experiment to investigate individuals' receptivity during four VR activities (Loading, 360 Video, Treasure Hunt, Rhythm Game) to message notifications delivered using three types of displays (head-mounted, controller, and movable panel). While higher engagement generally led to higher perceptions that notifications were ill-timed and/or disruptive, the suitability of notification displays to VR activities was influenced by the time-sensitiveness of VR content, overlapping use of modalities for delivering alerts, the display locations, and a requirement that the display be moved for notifications to be seen. Specific design suggestions are also provided.

Author Keywords

Virtual reality; notification systems; interruptibility; receptivity; eye-tracking

CCS Concepts

• **Human-centered computing~Laboratory experiments**; *Virtual reality*

INTRODUCTION

In recent years, a variety of immersive virtual reality (VR) applications have been developed and popularized. All leverage various modalities to provide their users with diverse experiences in immersive virtual environments. By their nature, these immersive experiences can result in users losing connectedness with the real world, including by dimming their awareness of incoming calls, text messages, and other phone notifications. Far from wanting to fully drop out of reality, however, VR users may desire to be notified about real-world events such as messages, and in some cases report feelings of being disconnected as causing them stress and anxiety [47,48]. Despite projections that users are likely to spend more time engaged in VR activities in the future, and thus to have even greater needs to stay informed of real-world events during VR sessions, our understanding of users' receptivity to real-world notifications within VR environments is limited, as is our knowledge of such notifications' potential disruptiveness in those environments.

Prior work has shown that external interruptions during tasks could make people feel disrupted [6, 12, 13] and negatively affect their emotions and overall experiences [34, 60]. These phenomena can be also influenced by task types [13] and how notifications are presented [38]. The current body of interruptibility and interruption research has mainly focused on desktop [14, 26] and mobile-platform contexts [9, 16, 46, 52]. Two recent studies explored notification presentation in VR [20, 59], but neither looked at how users' receptivity and recall of real-world notifications would be affected by specific VR activities and notification-display types.

To fill this research gap and to inform the design of future notification systems within VR platforms, we conducted mixed-methods research comprising a within-subjects experiment examining the effect of VR activity and notification display on users' perceived disruptiveness, perceived timeliness, and recall; and a semi-structured interview that explored their likes and dislikes across three kinds of notification display – head-mounted (HMD), controller, and movable panel – and four VR activities: Loading, 360 Video, Treasure Hunt, and Rhythm Game, each of which we presumed have different time-sensitiveness characteristics and visual-attention requirements. Our research questions were:

RQ1: In which VR activities do users perceive notifications as the most disruptive, and have the worst recall of them?

RQ2: Which notification displays are more suited to presenting notifications across various VR activities?

To the best of our knowledge, ours is the first study to investigate users' receptivity to real-world notifications via multiple notification displays across multiple VR activities. Its main contributions are, first, its finding that the more intensely the participants were engaged in a VR activity, the more likely they were to perceive notifications as disruptive and ill-timed, but that their recall of notifications fell into two distinct groups, depending on the suitability of the

CHI '20, April 25–30, 2020, Honolulu, HI, USA

ACM. ACM 978-1-4503-6708-0/20/04...$15.00
https://doi.org/10.1145/3313831.3376228

notification display. Its second main contribution is our finding that the time-sensitiveness of the VR activity, overlapping use of modalities for alerts, the display location, and requirements that users physically move to see alerts all influenced user perceptions' of notification displays' suitability. Based on these findings, we propose preliminary design recommendations for the mechanisms and appearance of future VR notification systems.

RELATED WORK

Interruptibility of Ongoing Task

Prior research by Bailey et al. suggested that interruptions have a disruptive effect on both a user's performance of an ongoing task and his/her emotional state [6]. Interruptions have been linked to annoyance and anxiousness [1, 6], and to a lower subjective sense of well-being [60]. Various other studies have explored the optimal timing for notification delivery. For instance, Bailey et al. [5] suggested that an attention-aware system that defers presenting peripheral information until coarse boundaries are reached during task execution could mitigate the negative impact of such information's arrival. Some studies have used system-usage features to predict opportune moments for notification [17–19], while others have investigated mobile activity [15, 24] and mental workload [28] for this purpose. Users have also been found to perceive varying levels of disruption depending on what tasks they are performing [12, 13], and that mobile interruptibility is influenced by their levels of task engagement [40, 45]. However, all such interruption research has been conducted on desktop or mobile platforms rather than in VR environments.

Presentation of Interruption

How interruptions are presented also affects users' receptivity [4, 38, 53], and Kreifeldt and McCarthy used different user-interface designs to reduce the negative effects of such interruptions [34]. Other researchers have modified notifications with the aim of increasing positive perceptions of interruptions [21, 55]. In a non-desktop, smart-home environment, Volt et al. [58] found that differences in the placement of notifications – e.g., on smartphones, next to the sending appliances, or on the user's body – influenced their perceived suitability. Some scholars have advocated peripheral displays for interruption management [23] and notification delivery [11, 36]. Maglio [37], for example, investigated how peripheral-display design could mitigate the negative impact of interruptions on users' performance of their primary tasks, while Costanza [11] suggested that such displays are more socially acceptable than some alternatives because of their unobtrusiveness. In the particular case of VR, some researchers have provided design guidelines [2, 29] and others have explored specific interactions [54], but two research attempts can be singled out as most relevant to the present work. The first was Ghosh et al.'s [20] empirical study of the noticeability and perception of five different VR interruption scenarios across six combinations of modalities. The same study also explored several ways of presenting notifications, including avatar display and ambient display. However, it only tested these designs in combination with one primary task, and did not investigate *how* they led to different disruption perceptions and recall across different VR activities. Second, Zenner [59] proposed the concept of adaptive, immersive ambient notification display, with the aim of leaving users in an unbroken state of VR immersion despite notifications appearing. But again, this research did not investigate users' receptivity to notifications presented via different modes of display in different VR activities. Accordingly, the current paper presents novel results regarding the main effects of VR activities and notification display, as well as the interaction effects between the two, on users' receptivity. Such findings highlight the need to consider the suitability of specific notification-display methods to specific VR activities.

THE EXPERIMENT

We conducted a within-subjects experiment with 40 participants, using the three types of notification display and four VR activities mentioned above, with the aim of learning how notifications' perceived disruptiveness, perceived timeliness, and recall differed across such activities, and whether any specific displays were more or less suited to presenting notifications during any specific VR activities.

VR Activities

The four VR activities we designed varied in two dimensions, *time sensitiveness* and *visual attention*, as explained below.

Waiting for system loading (hereafter, "Loading"; Figure 1, top left) was inspired by the Steam VR Home. Users are placed in a virtual room, which they can choose to explore or rearrange, or simply do nothing while waiting for the system to finish loading. As such, this VR activity was deemed non-time-sensitive and to require low visual attention.

Watching a 360 video ("360 Video"; Figure 1, top right) is a common VR activity for individuals wearing HMDs. We chose a 360 video of Antarctic nature exploration. The video was time-sensitive, and might occasionally take up more of users' visual attention if they found particular content interesting. Thus, we assumed it was relatively more visually engaging than Loading.

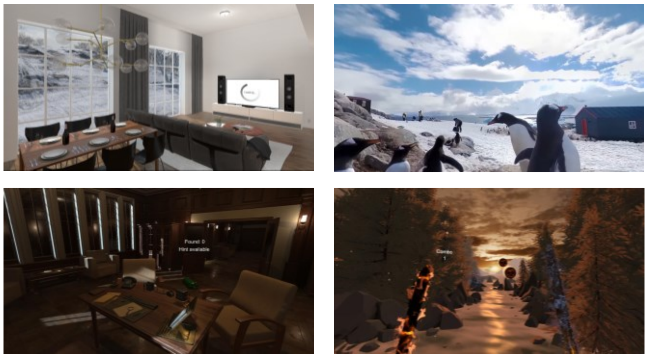

Figure 1. Waiting for system loading (top left), watching a 360 video (top right, © AirPano), playing a treasure hunt (bottom left) and playing a rhythm game (bottom right).

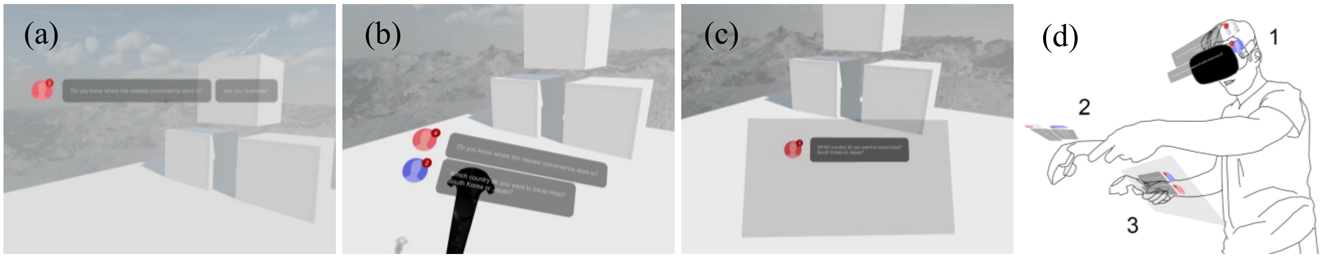

Figure 2. Three notification displays and three communication scenarios. HMD with Scenario 2 (a); controller display with Scenario 3 (b); movable panel display with Scenario 1 (c); and a sketch of all three notification displays (d), where HMD=1, controller=2, and movable panel=3.

Playing a treasure hunt ("Treasure Hunt"; Figure 1 bottom left) is a popular VR gaming format. In our game, the participant needed to locate three hidden objects by walking around an office and opening drawers. The game had no time limit, and they could ask for one hint every 60 seconds. The hints obtained and the number of remaining objects was shown on a floating panel. If all three objects were found before the last notification was sent, the game was extended into a bonus round. By its nature, this activity required a moderate level of visual attention, but was not time-sensitive.

Playing a rhythm game ("Rhythm Game"; Figure 1, bottom right) was inspired by a popular VR game called Beat Saber[1]. Our version allows participants to wave their controllers as swords, and slice balls of the same colors to garner points; and when a correct ball is hit, there is vibrating feedback from the controller. In our version, the score was displayed frontally. This activity was highly time-sensitive and required constant visual attention.

Notification Displays

The visual-notification displays we designed were inspired by prior research and common VR applications (Figure 2). Per the suggestion by NotifiVR [20] that haptic modality increases notifications' noticeability, we added vibration feedback to the controller when users received a notification, as in many existing VR applications. The notification appears at the top of the display pad and moves from left to right and from the top downward when new messages arrive, in line with most people's reading patterns [62].

The **HMD** (Figure 2a) was fixed to the upper left corner in the user's field of view: specifically, in the near-peripheral region, about 25 degrees from the line of sight [61].

The **controller display** (Figure 2b) was inspired by NotifiVR [20]. It showed notifications on a pad attached to the controller held in the user's non-dominant hand. Thus, the position of a notification could be changed via hand movements, much like on a smartwatch.

Lastly, the **movable panel display** (Figure 2c) was inspired by Facebook Space's[2] information pad. Our design consisted of a transparent black panel that showed notifications, initially placed approximately 0.3 meters in front of the user's abdomen (in keeping with the "touch UI zone" in [2]), but users could move it to anywhere they preferred at any time. However, if they lost track of its position during VR activities where they needed to turn frequently, the panel followed them horizontally to make itself easier to find.

Message Notification

We used instant messages as notifications in the experiment, as being the type of notification users most like to see most [51]. This was to reduce the likelihood that a participant would regard the notifications as disruptive simply because they were categorically uninterested in them. To add realism, we diversified message content, senders, and scenarios, but not to an extreme degree, since message content [16, 40, 57] and senders [35] have both been found to affect users' receptivity. We only included question-based messages, as they could reasonably be expected to arrive without any pre-determined conversational context. We designed two types of questions, factual-knowledge and opinion, on four topics – entertainment, shopping, places, and restaurants – adopted from [41]. We asked 10 volunteers to each generate 20 messages that comprised these types of questions and topics, and to word all of them in ways that suggested responses were not urgently required. Among the 200 resulting messages, we selected 48 that were distinct in terms of both topics and keywords, so that when recalling these notifications participants would not mistakenly recall because questions were similar. Next, we asked the participants to provide us with the names and profile photos of the three contacts they most frequently exchanged messages with, to serve as the ostensible sources of the messages they received during the experiment. In the final step, we established three communication scenarios, with the aim of rendering the messages' arrival more realistic to the participants (Figure 2). These were: one person sending a question message (Scenario 1); one person sending two messages successively (a greeting message followed by a question message, Scenario 2); and two persons, each sending a question message (Scenario 3).

Recruitment and Participants

The 40 participants consisted of 22 females and 18 males, aged 20-27 (M=23.3, SD=1.62); 34 (85%) were students. Five (12.5%) had never used a VR device before; 15 (37.5%) had used one fewer than times; and nine (22.5%) had used

---

[1] https://beatsaber.com/

[2] https://www.facebook.com/spaces

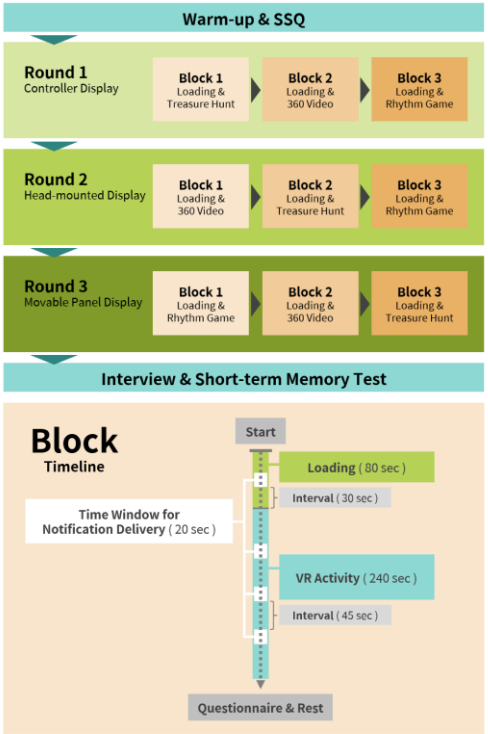

Figure 3. A sample of the study procedure (top) and the timeline of each block (bottom).

one 11-20 times. The remaining 11 (27.5%) had used VR more than 20 times. The participants were balanced in terms of their self-reported receptivity to message notifications, with 10 individuals fitting into each of the following four categories: 1) tend to ignore notifications; 2) tend not to deal with notifications immediately, but check and respond later; 3) tend to check notifications immediately, but respond later; and 4) tend to check and respond immediately.

Study Procedure

We used HTC Vive equipped with Tobbi Pro's eye-tracking hardware. The participants were informed of the study's goals, and then given a tutorial on HTC Vive. They then put on the headset and the over-the-ear headphones, and we performed inter-pupillary distance adjustment and eye-tracking calibration. Before the main phase of the experiment, the participants completed a simulator-sickness questionnaire (SSQ) [31] and performed a warm-up task in which they experienced all three notification displays and all four VR activities, to ensure that they understood how to operate the equipment and knew the rules of the two games.

The experiment *per se* consisted of three rounds (Figure 3, top). In each round, participants used a random combination of one designated notification display. Each round contained three blocks (Figure 3, bottom), in each of which participant experienced one of the three VR activities (360 Video, Treasure Hunt, and Rhythm Game) preceded by a Loading activity to simulate system loading. The order of the three blocks in each round was randomized. A VR activity was designed to be 240 seconds long (with Treasure Hunt being an exception because we set no time limit for it to let participants felt no time pressure during the activity). In each block, a notification first arrived in Loading, which was sent randomly within a 20-seconds time window between two 30-second intervals. Thus, each occurrence of Loading lasted 80 seconds, and three Loading also totaled 240 seconds. After Loading, three notifications arrived during the VR activity (or the first 240 seconds in Treasure Hunt). Similarly, each of the three notifications was sent within a 20-second time window, and a 45-second interval was placed between any two such time windows. Each notification lasted 12 seconds (including a one-second fade-in and a one-second fade-out).

After each block (Loading + a VR activity), participants took off the headset and answered a post-task questionnaire that contained four sections, covering: three questions from the SSQ (fatigue, vertigo, general discomfort level); a recall test regarding messages; perceptions of each messages; and the participant's level of engagement in the activity. In the perceptions section, we showed them screen recordings to help them recall the moments when they had received message notifications. The researcher also debriefed with the participants to obtain their feedback of using the displays within the VR activities. Lastly, each participant was given a short-term memory test to measure their general recall ability. The obtained comments served as guidance for the post-study interview after they completed all of the three rounds.

Outcome Variables and the Predictors

We measured four outcome variables: activity engagement, and the perceived disruptiveness, perceived timeliness, and recall of notifications. The first three were measured via the same seven-point Likert scale, with 1="*strongly disagree*" and 7="*strongly agree*". To measure recall, we asked the participants a series of multiple-choice questions regarding which notifications they thought they had seen during the VR activity. Each such question was followed by one correct answer, three wrong answers, and two additional options: "*Not sure*" and "*I did not see all of them*".

We also measured a number of predictor variables that we assumed could affect our outcome variables. For each notification, these included the perceived importance of the content; the order in which notifications arrived; the time elapsed between a notification's delivery and the end of the experiment (possibly influencing recall); and whether the participants had looked in the notification's direction. To measure the importance of message content, we asked whether the participant felt there was 1) no need to check and respond immediately, 2) a need to check immediately, but no need to respond immediately, or 3) a need to check and respond immediately. Gaze direction was obtained via Tobii Pro's eye-tracking device.

We also measured activity-level and participant-level predictors, including activity content, fatigue, vertigo, general discomfort levels, and general recall ability. Despite our attempts to control VR activity content across the three rounds, it did vary slightly so that participants would not feel bored due to seeing identical content repeated. For this

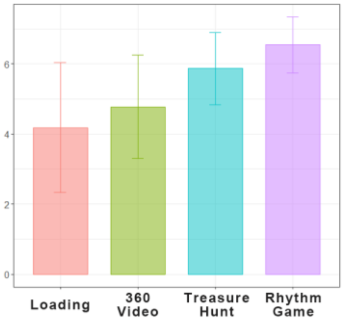

Figure 4. Means and SDs of engagement, by VR activity.

reason, VR activity content was also included as a predictor. Fatigue, vertigo and discomfort level were obtained from the SSQ, and their values subtracted from those measured at the beginning of the experiment. General recall ability was measured via a test adapted from Postman and Phillips [49], which showed 10 words on the screen, one every two seconds. The participants were asked to count down with the test for 15 seconds so that they would not concentrate on remembering these words, much as when seeing notifications in VR activities. The participants then completed 10 multiple-choice questions about which words had been shown to them.

DATA ANALYSIS

We built mixed-effects logistic regression models to examine the effects of the predictors on the outcome variables. We chose this statistical-analysis approach because each participant had 36 repeated observations, and we presumed that individual differences among participants would be large. Participant's code number was included as a random factor to account for individual difference. As some of the outcome variables were Likert-scaled responses, we used cumulative-link mixed models for these ordinal variables [63]. All the aforementioned predictors were included in the regression models, along with background data including gender, age, VR experience, and self-reported receptivity. However, score of the general recall ability were included only in the recall model. We built models based on our seven main-effect categories (e.g., four VR activities and three notification-display modes) as the reference level from which all pairwise comparisons between the categories were obtained. We also modeled a two-way interaction between activities and displays, and used the likelihood-ratio test [64] to examine their interaction effect in each model. Finally, based on an assumption that whether the participants had seen notifications would affect both their recall of such notifications and their perceptions of such notifications' disruptiveness, we created two datasets, one including all notifications, and the other including only those that eye-tracking data confirmed as having been seen.

We used affinity diagramming [30] to analyze the qualitative data. The themes that emerged through iterative grouping and labeling included preferences about notification-display mode, both when concentrating and when viewing time-sensitive content; and likes and dislikes about each mode.

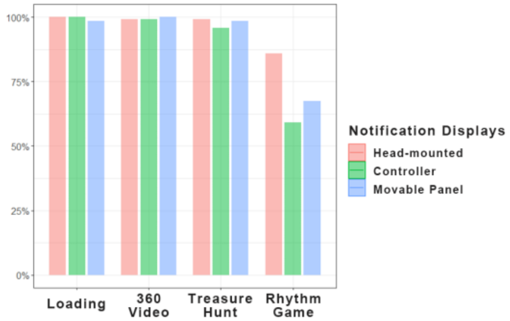

Figure 5. Rates of seen notifications, by VR activity.

RESULTS

We obtained 1,440 notification data points, each of which was associated with the predictors and outcome variables.

Engagement, and Notifications Actually Looked At

We first looked at participants' engagement and the percentages of notifications they had actually looked at across VR activities. As Figure 4 indicates, they were the most engaged in Rhythm Game (M=6.54), followed by Treasure Hunt (M=5.87), and 360 Video (M=4.78), and least engaged by Loading (M=4.18). All such differences were highly statistically significant (Rhythm Game vs. Treasure Hunt: Z=6.443, P<0.001; vs. 360 Video: Z=11.144, P<0.001; vs. Loading: 12.749, P<0.001; Treasure Hunt vs. 360 Video: Z=6.863, P<0.001 vs. Loading: Z=9.020, P<0.001; and 360 Video vs. Loading: Z=2.797, P=0.005). These results seem to imply that greater time-sensitiveness and higher demands on visual attention both led to higher engagement.

Figure 5, showing the percentages of notifications the participants actually looked at, indicates that they failed to see a significant portion of notifications in Rhythm Game, especially when alerts were sent via the controller (59.2%) or the movable panel display (67.5%). This suggests that in VR activities that required relatively high-level visual attention and time-sensitive, the participants easily missed notifications, especially when the display they were using was not fixed in a particular position.

Perceived Disruptiveness

The more engaged participants were in their VR activities, the more disruptive they perceived notifications to be. Thus, they perceived the highest amount of disruption during Rhythm Game (M=5.56) (Figure 6, top left), followed by Treasure Hunter (M=3.71), 360 Video (M=2.71), and Loading (M=1.61).

All differences among the VR activities were significant at the P<0.001 level. We did not see a main effect of notification display. However, likelihood-ratio testing revealed a marginal interaction effect between VR activity and notification display (LR stat=12.634, df=6, P=0.049). In 360 Video, the participants perceived the notifications presented via controller display as more disruptive (M = 3.09) than those presented via HMD (M=2.52, Z=-2.697, P=0.007) or via movable panel (M=2.51, Z=-3.143, P=0.002).

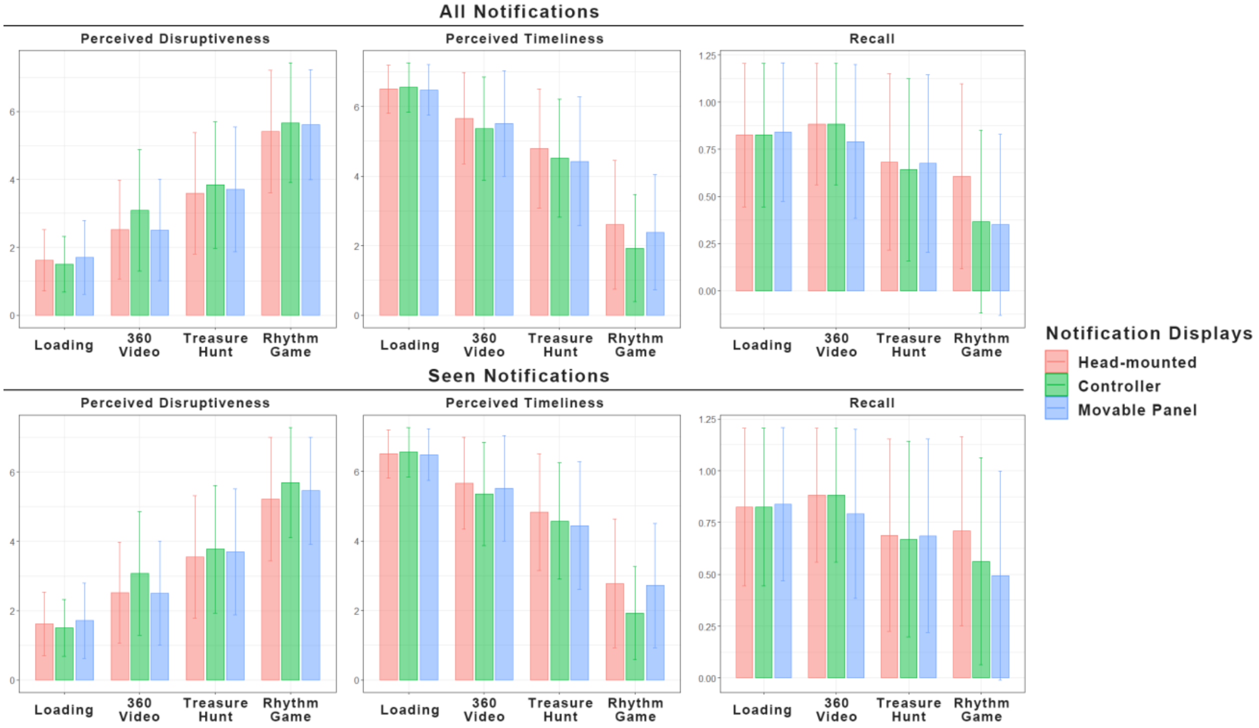

Figure 6. All notifications' means and SDs of the outcome variables: Perceived Disruptiveness (top left), Perceived Timeliness (top center) and Recall (top right); seen notifications' means and SDs of the outcome variables: Perceived Disruptiveness (bottom left), Perceived Timeliness (bottom center) and Recall (bottom right).

When we only considered those notifications that the participants had actually looked at (Figure 6, bottom left), all the previously noted differences held true, but in Rhythm Game, alerts presented via controller display (M=5.69) were perceived as significantly more disruptive than those presented via HMD (M=5.22, Z=-2.016, P=0.044) but not movable panel (M=5.47, Z=-1.476, P=0.140). This implies that some notifications in Rhythm Game were not perceived as disruptive simply because the participants were not aware of them; among those they did see, they reported the ones presented via controller display was more disruptive than the ones presented via HMD.

Timeliness of Visual Notifications

We observed a similar pattern in the perceived timeliness of notifications (Figure 6, top center): i.e., the more engaged the participants were in their current VR activity, the more likely they were to perceive notifications' timings as bad. Thus, they were seen as especially ill-timed in Rhythm Game (M=2.30, as compared to Loading: M=6.51; 360 Video: M=5.52; and Treasure Hunt: M=4.58). All such between-activity differences were significant at the P<0.001 level. Again, rather than any main effect of notification display, we observed a strong interaction effect of VR activity and display type (LR stat=18.102, df=6, P=0.006). In Rhythm Game, the participants perceived notifications presented via controller display as the worst-timed (M=1.93, vs. HMD: M=2.6, Z=-4.512, P<0.001; vs. movable panel: M=2.38, Z=4.121, P<0.001). When only viewed notifications were considered, these results were largely unchanged.

Recall of Notifications

The pattern of notification recall was somewhat different (Figure 6, top right). The participants could recall the notifications in 360 Video as well as in Loading (Z=0.704, P=0.482). However, their recall of notifications was significantly lower in Treasure Hunt and Rhythm Game (Loading vs. Treasure Hunt: Z=-3.385, P<0.001; Loading vs. Rhythm Game: Z=-4.188, P<0.001; 360 Video vs. Treasure Hunt: Z=-4.038, P<0.001; 360 Video vs. Rhythm Game: Z=-5.017, P<0.001; Treasure Hunt vs. Rhythm Game: Z=1.232, P=0.218). These results imply that in the two VR activities entailing more visual attention, people were less likely to recall the notifications they received. As before, however, we observed no main effect of notification display but again an interaction effect between it and VR activity (LR stat=19.105, df=6, P=0.004). In Rhythm Game, the participants could recall the notifications presented via HMD (M= 60.83%) much better than they recalled those presented via the controller display (M=36.67%, Z=-3.526, P<0.001) or via the movable panel (M=35%, Z=-4.313, P<0.001). In 360 Video, they had markedly worse recall of notifications presented via movable panel (M=79.17%) than of those presented via either HMD (M=88.33%, Z=2.408, P=0.016) or the controller (M=88.33%, Z=2.659, P=0.008).

As expected, in the dataset comprising only notifications that had been looked at (Figure 6, bottom right), we observed a significant increase in the recall of notifications for all displays (HMD: 75%=> 77.87%, controller: 67.92%=> 75.53%, movable panel: 66.46%=> 72.08%), suggesting that lack of recall was linked to a failure to observe alerts in the first place. However, even now considering viewed notifications, in Rhythm Game, the difference in notification recall between HMD (M=70.87%) and controller display (M=56.34%) was still significant (Z=-2.175, P=0.03), The difference between HMD (M=70.87%) and movable panel display (M=49.38%) also remained significant (Z=-3.162, P=0.002). Interestingly, movable panel display was associated with the lowest recall rates during time-sensitive VR activities (i.e., 360 Video and Rhythm Game) regardless of whether the notifications were seen. HMD, in contrast, was the most reliable display overall, as measured by recall rates across all VR activities.

Qualitative Findings

Our qualitative findings tended to support the quantitative results, as explained in detail below.

No Notifications during High Visual Attention

When the participants felt that an aspect of their current VR activity required high visual attention, such as searching for a hidden object in Treasure Hunt, visually tracking the balls in Rhythm Game, or viewing the most interesting parts of the 360 Video, they preferred not to receive any visual notifications, regardless of the display method. As P28 stated, "*The shot is zooming in on the seals! I found the notification more disruptive when I was looking at something I was more interested in.*" Visual notifications, in particular, were often in the way of what the user was trying to see: with some reporting that notifications next to the controller blocked their view of the working area when they needed to open or move something. And, while the position of the movable panel display was determined by the participants, they still sometimes had their lines of sight blocked by notifications when using it, because they could not anticipate where they would be looking. HMD was fixed to the upper left corner of their field of view, but some participants nevertheless complained that it was distracting: "*I considered myself pretty focused when finding the treasure, yet once the notification popped out, it would grab my attention immediately. So, it is quite disrupting to me*" (P18).

Some participants also said they disliked receiving vibration-based alerts when concentrating. Though this notification method did not block their vision, they found it caught them off guard and interrupted their engagement. As P31 put it, "*I was looking deep into the scene; therefore, I was scared when the controller suddenly vibrated!*"

Quick Access and Less Physical Effort when Viewing Time-Sensitive Content

In the time-sensitive VR content conditions, participants tended to find notifications disruptive if they felt they might miss content because of viewing them. This did not seemingly arise from a sense that their attention was being interrupted, but from a reluctance to miss content in the moment: "*[S]ince the video will continue to play, I might miss some content if I choose to read the notifications*" (P19). They found notifications less disruptive when the VR content was non-time-sensitive, as P20 explained: "*[In Treasure Hunt] I can take my time to find the object. I can pause the game without being punished for missing content. I can even reply to the message if I want, no big deal.*"

Given these preferences, when the VR content was time-sensitive, the participants liked designs that enabled them attend to notifications quickly and without much physical effort. Controller display and movable panel were thus particularly disfavored in such cases, as they had to physically move the display to deal with alerts, either by raising their hands in the case of the former (P31, P36), or by placing themselves in a 'bad position' vis-à-vis the main VR content, in the case of the latter (P22). This problem was especially serious in Rhythm Game, because, as P33 noted, "*In Rhythm Game, the disadvantage of Movable Panel is that when it is in a bad position, I don't have enough time to move it back.*"

Participants generally favored HMD when viewing time-sensitive content because it persistently showed notifications in the same place, thus saving them time. P21 said, "*The position is fixed, so you know where it is when you want to see it.*" Similarly, P39 said, "*Since it [HMD] is right in front of me, I can determine whether the text is important or not just with one glance in the middle of a task.*"

Likes and Dislikes regarding Notification Designs

Head-mounted Display

Nearly all positive comments about HMD involved its fixed location in the upper left corner, which not only made its notifications feel quicker to read, but also caused less visual blocking of the main VR content. However, some participants complained about their lack of freedom to place it in a more noticeable location when needed.

Controller Display

Controller display was considered an intuitive way of delivering notifications, in that the participants found it natural to look at the source of vibration, as they did with their smartphones ("*it is similar to checking phones in real life*" [P8]). Their major dislike of controller display, as noted briefly above, was that they could only see its notifications only when raised their hands. Despite finding this relatively non-interruptive, insofar as they would not see notifications' content so long as their hands remained down, it was nevertheless inconvenient, especially when viewing time-sensitive content, as mentioned earlier. In part this was because they would manually check whether they had just missed messages, as P29 explained: "*I won't see the message if I don't lift [the controller], so I am afraid that the important message will be missed.*" Another major negative aspect of controller display involved the multi-purpose role of the controller; i.e., participants found it troublesome to read notifications while also using the controller: "*Though it [i.e., the notification] being attached to the controller makes it easier to control where it appears, I can't really choose to place the controller wherever I wanted in a Rhythm Game.*

In that case, attaching notifications to the controller is instead a constraint because I can't flexibly adjust the position of the notification." (P18). That notifications attached to the controller blocked content when the controller was in use and that participants could not clearly read notifications while they were moving the controller also helps explain why users perceived its notifications as the worst-timed, and felt alerts were most disruptive when using the controller display in Rhythm Game.

Movable Panel

The freedom to move the notification panel around, e.g., to a place where it was easy to read the notifications, was the primary feature participants liked about this display. However, they disliked that they did not necessarily know where such 'good' locations were; and many mentioned accidentally dragging the display to an angle at which its notifications were too blurry to read. Some suggested that the system should provide recommended positions for the panel to prevent this from happening. Other negative comments included the panel's large size (even when no notification had arrived) and its non-user-controllable opacity.

Non-visual Notification: Vibrations

The participants also made direct comparisons between their experiences of haptic feedback and visual alerts. They agreed that vibrations were more intrusive and thus harder to miss, they disagreed on whether this was a positive or negative characteristic. In particular, when the VR activity itself *also* delivered haptic feedback, as in Rhythm Game, the participants had quite divergent reactions toward their experience of two functions sharing the same modality. Some complained that they could not distinguish message vibrations from game ones, and thus missed the former, but others preferred not to notice message notifications, as this made them feel less interrupted. Still others complained that their misjudgments about the sources of vibrations could be quite annoying, in that they felt they needed to check manually whether there were incoming notifications. One noted that he lost many "ball hits" in Rhythm Game when the controller vibrated unexpectedly: "*The game uses vibration, so when the notification came [... the feedback] did not match what I saw, so I missed the next ball*" (P11).

DISCUSSION

When VR platforms become common features of people's daily lives, they will inevitably need to present real-world notifications, due to users' persistent tendency to stay connected with their social worlds (e.g., [48]). Our findings suggest that informing users about real-world events such as message arrival within VR can be acceptable, but that users' receptiveness to such notifications can vary enormously, depending on what VR activity they are engaged in. Also, the more engaged our experimental participants were in a VR activity, the more disruptive and ill-timed they perceived real-world notifications to be. This is consistent with the findings of previous interruptibility studies conducted on desktop devices [12, 13] and mobile platforms [40, 45].

However, our participants' recall of notifications seemed to be influenced not only by the intensity of their VR engagement but by other factors. For instance, low recall could be attributed either to participants failing to see notifications at all, or seeing them imperfectly because the display was poorly positioned. These problems were particularly acute when the VR content was time-sensitive. All notifications in the experiment were visible for 12 seconds each, but participants often missed them not because they had failed to perceive them, but rather because they made the active choice to attend to VR content instead, out of a desire not to miss the latter. The lower recall for notifications presented via the movable panel display than via the other displays, in both 360 Video and Rhythm Game, seemed to exemplify this negative influence of time-sensitiveness on users' willingness to move the panel to a better position. This effect was *not* observed in the two presumed non-time-sensitive VR activities.

In cases where the user simply did not know that a notification had arrived, on the other hand, a further contributing factor was the shared use of controller vibration between Rhythm Game and the notification system. That is, since they were constantly receiving vibration feedback as part of the game, users typically could not tell whether a given episode of vibration came from the notification system or not. This led to a much higher proportion of notifications being missed in Rhythm Game than in the other tested VR activities. HMD display, in contrast, appeared highly suitable for notification delivery when the VR content was relatively time-sensitive, because the fixed location of its alerts was easy to remember and easy to access.

Our findings regarding the disadvantages of the controller display also support previous suggestions [20] that such displays are ill-suited to VR tasks requiring numerous hand actions. In Rhythm Game, which required extensive hand movements to play, notifications presented via controller display – if seen at all – were perceived as the worst-timed of all visual notifications in our experiment, as well as the most disruptive, because they blocked the viewing area and did not allow participants to see notifications clearly while the controller was moving. In addition, we showed that controller display led to stronger perceptions of disruptiveness and poor timing than other displays did in 360 Video, not because it was multi-purpose but because seeing its notifications required additional effort from users.

Thus, returning now to our research questions, the VR activity in which users generally perceived notifications as the most disruptive and recalled the least well was Rhythm Game. However, regarding the relationships between notification displays and VR activities, we found that the impact of VR activity on notification receptivity is not simply about how much engagement it demands from users, but also about the interplay between the time-sensitiveness of its content and the characteristics of the notification-display. In particular, our results imply that the key factors in the perceived suitability of notification-delivery methods were: 1) the time-sensitiveness of the VR content, 2) the use of the same modality for message notifications and VR content, 3) the location of the display, and 4) the need to physically move to see notifications.

Finally, it was interesting that many participants preferred to be notified in the same ways as on their desktop or mobile platforms, including in terms of placement within their field of view; and familiarity with mobile phones was a major reason participants liked controller display. However, it is likely that such preferences based on non-VR technologies may gradually diminish as VR becomes more prevalent.

DESIGN RECOMMENDATIONS

Prior researchers have attempted to detect break-points or opportune moments within individuals' use of engaging desktop applications, as a means of making them more receptive to notifications [14, 18, 26]. Based on our findings, we argue that future notification systems on VR platforms ought to be attention-, content-, and modality-aware. Specific design guidelines are proposed below.

The VR Platform Should Be Attention-aware

When users are highly engaged in VR activity, the VR platform should not inform them of notifications, except those that are especially urgent/important. Given the ready availability of VR headsets equipped with eye-tracking functionality (e.g., HTC VIVE Pro Eye), we recommend that VR platforms use eye-tracking to detect users' concentration and engagement levels, as a means of determining appropriate break-points at which to deliver notifications [25, 42, 56]. In any modality, notifications delivered when users are highly engaged are very likely to be regarded as highly disruptive, thereby harming their immersion in the VR activity and thus their overall experience; moreover, such notifications are less likely to be recalled clearly. As an alternative, a small icon placed in the user's peripheral vision, carrying updates on the number of as-yet unviewed notifications, might be more acceptable than displaying full notification content. Then, when the VR activity becomes less attention-demanding, the VR platform could provide some further reminder about these unread notifications.

The VR Platform Should Be Content-aware

Based on our data, we recommend that VR platforms maintain awareness of at least two aspects of the content their users are currently accessing: 1) its time-sensitiveness, and 2) the key region of interest/importance. With regard to the first, when the platform is aware that the content is time-sensitive, it should display real-world notifications spontaneously at a fixed location in the user's peripheral vision, based on their current gaze direction, without requiring them to manually call up notifications to see them. This will help ensure that they do not feel forced to decide between enjoying VR content and responding to real-world events. Though our data provides no direct evidence regarding its effectiveness, such an approach has been found effective in prior research [10, 33] and could help address 'blurry notifications' caused by bad angles.

Our participants generally disliked their region of focus being blocked by notifications: they reported the major drawback of HMD to be its inflexible location, which always blocked a specific region, and criticized controller-display notifications as blocking the working area. Therefore, we recommend that VR platforms allow their users to choose where notifications appear, from among a predetermined set of peripheral locations. If the platform considers presenting real-world events to be worth the hassle of setting a policy for its app producers, it could also reasonably ask for metadata about the fixed locations where each app places its own important content, such as settings, profiles, or in-app notifications. Alternatively, the platform could adopt a more technical approach, detecting the region of interest/ importance dynamically, since that region might change from moment to moment. One version of such an approach might be to use eye-tracking data to detect regions that users frequently look at, attend to, and act upon [22, 32]. Another might be to detect the region using computer vision [27], or based on the rendering of the graphical elements visible in the screen. According to the detection outcome, the platform could then dynamically adjust the location so that it does not block major content. At a minimum, the platform should adjust notifications' locations when interference occurs frequently. On the other hand, when the user determines a location at which frequent interference is expected to occur by the platform, the platform should inform the user about this and recommend other regions. Although maintaining notification-display consistency is regarded as an important element of usability [43], different applications may position important and frequently occurring items in different places; thus, we believe flexibility and dynamic placement could have wider benefits that outweigh the potential harm to the user experience caused by notifications blocking the major VR content. Moreover, we believe such inconsistencies can be addressed and understood by an intelligent system, which can explain them clearly and intelligibly [3].

The VR Platform Should Be Modality-aware

Although HMD enabled our participants to miss fewer notifications, it mainly did so when the problem of missed notifications was caused by the overlapping use of vibration by both the notification display and Rhythm Game. Therefore, as NotifiVR [20] also suggested, we recommend that VR platforms' notification systems distinguish clearly among notification-presentation modalities, so that they can avoid the confusion that is often caused by using the same modality both for VR-activity feedback and real-world notifications. If an overlapping use of same modality is identified before the VR activity starts, the platform could inform the user about it and ask how he/she wants to be notified, with options including other modalities and not being notified at all. Recent research has shown that users may be able to identify the source of a notification based on alert patterns [8]; and different patterns of haptic feedback used in VR have also been explored extensively in HCI research [7, 44]. Nevertheless, VR platforms will then need to dynamically determine a pattern distinct enough from those used in the current VR activity: a task that will be progressively more challenging as more patterns are used. Creating a stand-alone wearable device, such as a ring, for sending haptic notifications [50] might also be viable.

When Sending Important/Urgent Messages

Finally, we regard it as important to heighten users' awareness of urgent/important notifications even when they are immersed in VR, as the cost of missing such notifications potentially outweighs the disruption caused. Though the HMD used in our study was relatively more noticeable than the other displays, it was still associated with poor recall in both Treasure Hunt and Rhythm Game. Therefore, platforms should try to identify even more obtrusive methods to use with urgent notifications, such as auditory alerts [20]. Also, because recent research has shown smartphone users are more receptive to notifications from certain sources [35] and that they rate as having high content importance [57], we envision a smartphone service that can infer the importance of notifications using these techniques and decide whether to deliver the notifications to the VR platform or to other devices in the multi-device environment [39]. Alternatively, VR systems themselves could offer customizable notification permissions, to allow important notifications (e.g., from a specific person/app, or containing specific keywords) to be presented in VR, but block others.

LIMITATIONS

Our experiments took place in a laboratory, so that numerous impacts on the participants' receptivity to messages within VR could be controlled. However, it is impossible to simulate the actual conditions of real-world VR use, as people's receptivity is highly situated and content- and source-dependent. VR users are likely to encounter a greater diversity of messages, display methods, and VR activities than we could possibly have provided. Also, there are numerous alternatives to the three notification displays and four VR activities we studied. For instance, ambient display could present notifications [59]; wearable devices other than controllers could be used to provide haptic feedback; and audio modalities such as ringtones and voice assistants could be incorporated among the set of non-visual notifications.

We also set an arbitrary notification duration of 12 seconds and did not explore alternative durations. In our opinion, 12 seconds was ample as a test of users' receptivity to a notification at a given moment. However, our results – particularly on the relation of missed notifications to the time-sensitiveness of VR activities – would inevitably have differed somewhat if we had set a longer or shorter notification duration, or kept notifications visible until participants took explicit action pertaining to them. In any case, lengthening notifications seems likely to cause them to 'stack' in users' fields of view, making them more noticeable perhaps, but also more likely to be perceived as disruptive. Therefore, the exact relationships to our outcome variables of notification duration, as well as notification fade-out approaches, merit further study.

Likewise, we did not explore notifications in other VR activities, which our participants might have had different perceptions of. However, because the typical duration of our experimental session was two and a half hours, exploring many factors would likely have caused user fatigue, impacting not only the accuracy of the results but also the participants' well-being. Moreover, our claim of the different time-sensitiveness of the VR activities are reasonable yet not validated assumption. We did not measure participants' perceived time-sensitiveness of the four VR activities, but such perception could vary from users to users. We regard our experiment as a preliminary study, and hope that other researchers will build upon its results to help VR practitioners design more user-friendly notification systems for their VR platforms.

CONCLUSION

As VR applications become more prevalent, and VR hardware improves such that longer usage can be endured, we anticipate that people will spend more time in VR and that VR devices will become full members of cross-device notification ecology. Thus, it is important to understand users' receptivity to real-world notifications during various VR activities, and to design the display of such notifications effectively. While our experiment only explored three notification displays, our aim was to understand the pros and cons of each one. We found that VR activity type had a main effect on users' perceptions of notifications, and that there was an interaction effect of VR activities and display designs on such perceptions. More importantly, we found what while engagement level was an important factor in whether users wanted to be notified, the time-sensitiveness of VR content, use of the same modality for message notifications and VR content, the location of the display, and the need to move the display to see notifications also influenced the perceived suitability of a specific notification display type. Thus, we recommend that designers of systems for VR platforms that inform users of real-world events such as incoming messages concentrate on content- and modality-awareness. While we hope that our findings and design recommendations will help boost HCI researchers' and VR practitioners' understanding of VR users' receptivity to real-world notifications, we freely concede that this research is only a start, and that further exploration the design space of notification systems in VR will be needed if we are to effectively bridge the gap between the virtual and real worlds.

ACKNOWLEDGMENTS

This research was supported in part by the Ministry of Science and Technology, R.O.C. (MOST 106-2221-E-009-141), and access to equipment was provided by the Communication and Cognition Lab, National Chiao Tung University. Special thanks to Dr. Li-Wei Chan, Dr. Wen-Chieh Lin and Dr. Da-Yuan Huang for their helpful feedback.

REFERENCES

[1] Piotr D. Adamczyk and Brian P. Bailey. 2004. If not now, when?: the effects of interruption at different moments within task execution. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 271–278.

[2] Michael Alger. 2015. Visual Design Methods for Virtual Reality.

[3] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (CHI ’19), 3:1–3:13. https://doi.org/10.1145/3290605.3300233

[4] E. Arroyo, T. Selker, and A. Stouffs. 2002. Interruptions as multimodal outputs: which are the less disruptive? In *Proceedings. Fourth IEEE International Conference on Multimodal Interfaces*, 479–482. https://doi.org/10.1109/ICMI.2002.1167043

[5] Brian P. Bailey and Joseph A. Konstan. 2006. On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. *Computers in Human Behavior* 22, 4: 685–708. https://doi.org/10.1016/j.chb.2005.12.009

[6] Brian P. Bailey, Joseph A. Konstan, and John V. Carlis. 2001. The effects of interruptions on task performance, annoyance, and anxiety in the user interface. In *In: Proceedings of IFIP TC.13 International Conference on Human–- Computer Interaction. IOS*, 593–601.

[7] Hong-Yu Chang, Wen-Jie Tseng, Chia-En Tsai, Hsin-Yu Chen, Roshan Lalintha Peiris, and Liwei Chan. 2018. FacePush: Introducing Normal Force on Face with Head-Mounted Displays. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (UIST ’18), 927–935. https://doi.org/10.1145/3242587.3242588

[8] Yung-Ju Chang, Yi-Ju Chung, Yi-Hao Shih, Hsiu-Chi Chang, and Tzu-Hao Lin. 2017. What Do Smartphone Users Do when They Sense Phone Notifications? In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers* (UbiComp ’17), 904–909. https://doi.org/10.1145/3123024.3124557

[9] Yung-Ju Chang and John C. Tang. 2015. Investigating Mobile Users’ Ringer Mode Usage and Attentiveness and Responsiveness to Communication. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (MobileHCI ’15), 6–15. https://doi.org/10.1145/2785830.2785852

[10] Çağla Çığ Karaman and Tevfik Metin Sezgin. 2018. Gaze-based predictive user interfaces: Visualizing user intentions in the presence of uncertainty. *International Journal of Human-Computer Studies* 111: 78–91. https://doi.org/10.1016/j.ijhcs.2017.11.005

[11] Enrico Costanza, Samuel A. Inverso, Elan Pavlov, Rebecca Allen, and Pattie Maes. 2006. Eye-q: Eyeglass Peripheral Display for Subtle Intimate Notifications. In *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services* (MobileHCI ’06), 211–218. https://doi.org/10.1145/1152215.1152261

[12] Edward Cutrell, Mary Czerwinski, and Eric Horvitz. 2001. Notification, Disruption, and Memory: Effects of Messaging Interruptions on Memory and Performance. 263–269.

[13] Mary Czerwinski, Edward Cutrell, and Eric Horvitz. 2000. *Instant Messaging and Interruption: Influence of Task Type on Performance*.

[14] Mary Czerwinski, Edward Cutrell, and Eric Horvitz. 2000. Instant messaging: Effects of relevance and timing. In *People and computers XIV: Proceedings of HCI*, 71–76.

[15] Joel E. Fischer, Chris Greenhalgh, and Steve Benford. 2011. Investigating Episodes of Mobile Phone Activity As Indicators of Opportune Moments to Deliver Notifications. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services* (MobileHCI ’11), 181–190. https://doi.org/10.1145/2037373.2037402

[16] Joel E. Fischer, Nick Yee, Victoria Bellotti, Nathan Good, Steve Benford, and Chris Greenhalgh. 2010. Effects of Content and Time of Delivery on Receptivity to Mobile Interruptions. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services* (MobileHCI ’10), 103–112. https://doi.org/10.1145/1851600.1851620

[17] James Fogarty, Scott E. Hudson, Christopher G. Atkeson, Daniel Avrahami, Jodi Forlizzi, Sara Kiesler, Johnny C. Lee, and Jie Yang. 2005. Predicting human interruptibility with sensors. *ACM Transactions on Computer-Human Interaction (TOCHI)* 12, 1: 119–146.

[18] James Fogarty, Scott E. Hudson, and Jennifer Lai. 2004. Examining the Robustness of Sensor-based Statistical Models of Human Interruptibility. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI ’04), 207–214. https://doi.org/10.1145/985692.985719

[19] James Fogarty, Andrew J. Ko, Htet Htet Aung, Elspeth Golden, Karen P. Tang, and Scott E. Hudson. 2005. Examining task engagement in sensor-based statistical models of human interruptibility. In *Proceedings of the*

SIGCHI conference on Human factors in computing systems - CHI '05, 331. https://doi.org/10.1145/1054972.1055018

[20] S. Ghosh, L. Winston, N. Panchal, P. Kimura-Thollander, J. Hotnog, D. Cheong, G. Reyes, and G. D. Abowd. 2018. NotifiVR: Exploring Interruptions and Notifications in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* 24, 4: 1447–1456. https://doi.org/10.1109/TVCG.2018.2793698

[21] Jennifer Gluck, Andrea Bunt, and Joanna McGrenere. 2007. Matching attentional draw with utility in interruption. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 41–50.

[22] Joseph H. Goldberg, Mark J. Stimson, Marion Lewenstein, Neil Scott, and Anna M. Wichansky. 2002. Eye Tracking in Web Search Tasks: Design Implications. In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications* (ETRA '02), 51–58. https://doi.org/10.1145/507072.507082

[23] Shameem Hameed, Thomas Ferris, Swapnaa Jayaraman, and Nadine Sarter. 2009. Using Informative Peripheral Visual and Tactile Cues to Support Task and Interruption Management. *Human Factors* 51, 2: 126–135. https://doi.org/10.1177/0018720809336434

[24] Joyce Ho and Stephen S. Intille. 2005. Using context-aware computing to reduce the perceived burden of interruptions from mobile devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 909–918.

[25] Bert Hoeks and Willem J. M. Levelt. 1993. Pupillary dilation as a measure of attention: a quantitative system analysis. *Behavior Research Methods, Instruments, & Computers* 25, 1: 16–26. https://doi.org/10.3758/BF03204445

[26] Eric Horvitz, Johnson Apacible, and Muru Subramani. 2005. Balancing Awareness and Interruption: Investigation of Notification Deferral Policies. In *User Modeling 2005*, Liliana Ardissono, Paul Brna and Antonija Mitrovic (eds.). Springer Berlin Heidelberg, 433–437. https://doi.org/10.1007/11527886_59

[27] Hou-Ning Hu, Yen-Chen Lin, Ming-Yu Liu, Hsien-Tzu Cheng, Yung-Ju Chang, and Min Sun. 2017. Deep 360 Pilot: Learning a Deep Agent for Piloting through 360° Sports Videos. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1396–1405. https://doi.org/10.1109/CVPR.2017.153

[28] Shamsi T. Iqbal and Brian P. Bailey. 2005. Investigating the Effectiveness of Mental Workload As a Predictor of Opportune Moments for Interruption. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '05), 1489–1492. https://doi.org/10.1145/1056808.1056948

[29] Jason Jerald. 2015. *The VR Book: Human-Centered Design for Virtual Reality*. Morgan & Claypool.

[30] Jiro Kawakita. 1982. The Original KJ Method.

[31] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology* 3, 3: 203–220. https://doi.org/10.1207/s15327108ijap0303_3

[32] Peter Kiefer, Ioannis Giannopoulos, Dominik Kremer, Christoph Schlieder, and Martin Raubal. 2014. Starting to Get Bored: An Outdoor Eye Tracking Study of Tourists Exploring a City Panorama. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (ETRA '14), 315–318. https://doi.org/10.1145/2578153.2578216

[33] Michaela Klauck, Yusuke Sugano, and Andreas Bulling. 2017. Noticeable or Distractive?: A Design Space for Gaze-Contingent User Interface Notifications. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (CHI EA '17), 1779–1786. https://doi.org/10.1145/3027063.3053085

[34] John G. Kreifeldt and M. E. McCarthy. 1981. Interruption as a test of the user-computer interface.

[35] Hao-Ping Lee, Kuan-Yin Chen, Chih-Heng Lin, Chia-Yu Chen, Yu-Lin Chung, Yung-Ju Chang, and Chien-Ru Sun. 2019. Does Who Matter?: Studying the Impact of Relationship Characteristics on Receptivity to Mobile IM Messages. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (CHI '19), 526:1–526:12. https://doi.org/10.1145/3290605.3300756

[36] 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* (ACE '14), 17:1–17:10. https://doi.org/10.1145/2663806.2663824

[37] Paul P. Maglio and Christopher S. Campbell. 2000. Tradeoffs in Displaying Peripheral Information. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '00), 241–248. https://doi.org/10.1145/332040.332438

[38] Daniel McFarlane. 2002. Comparison of four primary methods for coordinating the interruption of people in human-computer interaction. *Human-Computer Interaction* 17, 1: 63–139.

[39] Abhinav Mehrotra, Robert Hendley, and Mirco Musolesi. 2019. NotifyMeHere: Intelligent Notification Delivery in Multi-Device Environments. In *Proceedings of the 2019 Conference on Human*

Information Interaction and Retrieval (CHIIR ’19), 103–111. https://doi.org/10.1145/3295750.3298932

[40] 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* (CHI ’16), 1021–1032. https://doi.org/10.1145/2858036.2858566

[41] Meredith Ringel Morris, Jaime Teevan, and Katrina Panovich. 2010. What do people ask their social networks, and why?: a survey study of status message q&a behavior. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1739–1748.

[42] Minoru Nakayama, Koji Takahashi, and Yasutaka Shimizu. 2002. The Act of Task Difficulty and Eye-movement Frequency for the “Oculo-motor Indices.” In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications* (ETRA ’02), 37–42. https://doi.org/10.1145/507072.507080

[43] Jakob Nielsen and Rolf Molich. 1990. Heuristic Evaluation of User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI ’90), 249–256. https://doi.org/10.1145/97243.97281

[44] Roshan Lalintha Peiris, Wei Peng, Zikun Chen, Liwei Chan, and Kouta Minamizawa. 2017. ThermoVR: Exploring Integrated Thermal Haptic Feedback with Head Mounted Displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI ’17), 5452–5456. https://doi.org/10.1145/3025453.3025824

[45] Veljko Pejovic, Mirco Musolesi, and Abhinav Mehrotra. 2015. Investigating The Role of Task Engagement in Mobile Interruptibility. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (MobileHCI ’15), 1100–1105. https://doi.org/10.1145/2786567.2794336

[46] Martin Pielot, Rodrigo de Oliveira, Haewoon Kwak, and Nuria Oliver. 2014. Didn’t you see my message?: predicting attentiveness to mobile instant messages. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, 3319–3328.

[47] Martin Pielot and Luz Rello. 2015. The Do Not Disturb Challenge: A Day Without Notifications. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (CHI EA ’15), 1761–1766. https://doi.org/10.1145/2702613.2732704

[48] Martin Pielot and Luz Rello. 2017. Productive, Anxious, Lonely: 24 Hours Without Push Notifications. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (MobileHCI ’17), 11:1–11:11. https://doi.org/10.1145/3098279.3098526

[49] Leo Postman and Laura W. Phillips. 1965. Short-term temporal changes in free recall. *The Quarterly Journal of Experimental Psychology* 17, 2: 132–138. https://doi.org/10.1080/17470216508416422

[50] Thijs Roumen, Simon T. Perrault, and Shengdong Zhao. 2015. NotiRing: A Comparative Study of Notification Channels for Wearable Interactive Rings. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI ’15), 2497–2500. https://doi.org/10.1145/2702123.2702350

[51] Alireza Sahami Shirazi, Niels Henze, Tilman Dingler, Martin Pielot, Dominik Weber, and Albrecht Schmidt. 2014. Large-scale Assessment of Mobile Notifications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI ’14), 3055–3064. https://doi.org/10.1145/2556288.2557189

[52] Florian Schulze and Georg Groh. 2014. Studying How Character of Conversation Affects Personal Receptivity to Mobile Notifications. In *Proceedings of the Extended Abstracts of the 32Nd Annual ACM Conference on Human Factors in Computing Systems* (CHI EA ’14), 1729–1734. https://doi.org/10.1145/2559206.2581320

[53] Cheri Speier, Joseph Valacich, and Iris Vessey. 1997. The effects of task interruption and information presentation on individual decision making. 21–36.

[54] Hemant Bhaskar Surale, Aakar Gupta, Mark Hancock, and Daniel Vogel. 2019. TabletInVR: Exploring the Design Space for Using a Multi-Touch Tablet in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (CHI ’19), 13:1–13:13. https://doi.org/10.1145/3290605.3300243

[55] Dan Tasse, Anupriya Ankolekar, and Joshua Hailpern. 2016. Getting Users’ Attention in Web Apps in Likable, Minimally Annoying Ways. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI ’16), 3324–3334. https://doi.org/10.1145/2858036.2858174

[56] Y. Tateyama, Y. Matsumoto, and S. Kagami. 2004. Concentration detection by eye movements: towards supporting a human. In *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583)*, 1544–1548 vol.2. https://doi.org/10.1109/ICSMC.2004.1399851

[57] Aku Visuri, Niels van Berkel, Tadashi Okoshi, Jorge Goncalves, and Vassilis Kostakos. 2019. Understanding smartphone notifications’ user interactions and content importance. *International Journal of Human-Computer Studies* 128: 72–85. https://doi.org/10.1016/j.ijhcs.2019.03.001

[58] Alexandra Voit, Tonja Machulla, Dominik Weber, Valentin Schwind, Stefan Schneegass, and Niels Henze. 2016. Exploring notifications in smart home environments. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct - MobileHCI '16*, 942–947. https://doi.org/10.1145/2957265.2962661

[59] André Zenner, Marco Speicher, Sören Klingner, Donald Degraen, Florian Daiber, and Antonio Krüger. 2018. Immersive Notification Framework: Adaptive & Plausible Notifications in Virtual Reality. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI EA '18), LBW609:1–LBW609:6. https://doi.org/10.1145/3170427.3188505

[60] Fred RH Zijlstra, Robert A. Roe, Anna B. Leonora, and Irene Krediet. 1999. Temporal factors in mental work: Effects of interrupted activities. *Journal of Occupational and Organizational Psychology* 72, 2: 163–185.

[61] 2019. Peripheral vision. *Wikipedia*. Retrieved August 18, 2019 from https://en.wikipedia.org/w/index.php?title=Peripheral_vision&oldid=906668484

[62] F-Shaped Pattern For Reading Web Content (original eyetracking research). *Nielsen Norman Group*. Retrieved September 20, 2019 from https://www.nngroup.com/articles/f-shaped-pattern-reading-web-content-discovered/

[63] clmm function | R Documentation. Retrieved August 6, 2019 from https://www.rdocumentation.org/packages/ordinal/versions/2019.4-25/topics/clmm

[64] anova function | R Documentation. Retrieved August 21, 2019 from https://www.rdocumentation.org/packages/ltm/versions/1.1-1/topics/anova