NotifiVR: Exploring Interruptions and Notifications in Virtual Reality

Sarthak Ghosh, Lauren Winston, Nishant Panchal, Philippe Kimura-Thollander
Jeff Hotnog, Douglas Cheong, Gabriel Reyes, and Gregory D. Abowd

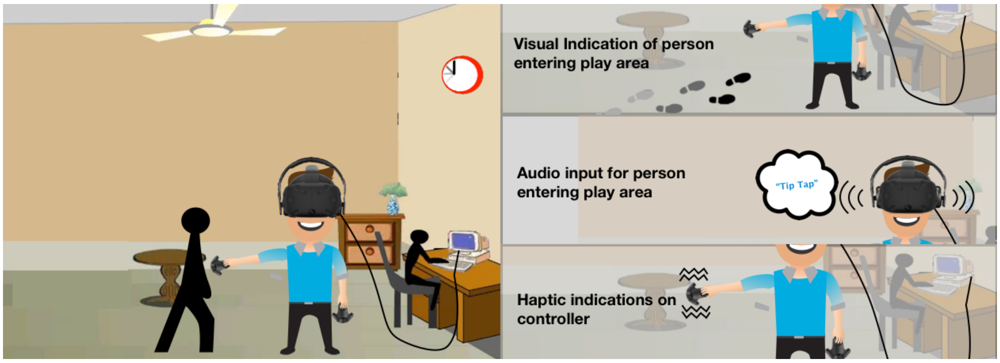

Fig. 1. *Left*: A person is immersed in VR and another person walks into the play area. *Right*: The user is notified of the event either visually (e.g., footprints), and/or through audio (e.g., localized "tip-tap"), and/or through haptics (e.g., controller vibrations).

Abstract—The proliferation of high resolution and affordable virtual reality (VR) headsets is quickly making room-scale VR experiences available in our homes. Most VR experiences strive to achieve complete immersion by creating a disconnect from the real world. However, due to the lack of a standardized notification management system and minimal context awareness in VR, an immersed user may face certain situations such as missing an important phone call (digital scenario), tripping over wandering pets (physical scenario), or losing track of time (temporal scenario). In this paper, we present the results of 1) a survey across 61 VR users to understand common interruptions and scenarios that would benefit from some form of notifications; 2) a design exercise with VR professionals to explore possible notification methods; and 3) an empirical study on the noticeability and perception of 5 different VR interruption scenarios across 6 modality combinations (e.g., audio, visual, haptic, audio + haptic, visual + haptic, and audio + visual) implemented in Unity and presented using the HTC Vive headset. Finally, we combine key learnings from each of these steps along with participant feedback to present a set of observations and recommendations for notification design in VR.

Index Terms—Virtual Reality, Notifications, Interruptibility, Multi-Modal, Feedback, Context Awareness

1 Introduction

Many virtual reality (VR) experiences are designed to immerse people in alternate worlds. The most convincing user experience is achieved by enhancing the sense of presence and enabling intuitive interactions. As such, VR experiences often engage 3 of our 5 major senses - visual, auditory, and tactile. Visual stimulation (through the display), auditory effects (through headphones), and tactile sensations (often through controllers), all play critical roles in turning the virtual "world" into a virtual "reality" [40]. However, captivating the senses creates a significant disconnect from the real world which is sometimes undesirable.

It is well known that wearing a head-mounted display (HMD) causes people to lose sense of their surroundings, which often leads to minor accidents like colliding with walls/furniture [28, 29]. Popular HMDs such as the HTC Vive provide such cautionary messages in their safety guide. A common solution is to designate an empty play-area for VR. However, this limits VR from being used in uncontrolled settings. It may also be difficult to cordon off a big enough VR designated area in a living room. Due to complete sensory immersion, users may also miss calls, text messages, and other notifications generated from their digital ecosystems. This disconnect will likely be even more apparent as VR enters the workplace. In addition, due to the lack of natural zeitgebers and extreme engagement, immersive environments may affect our perception of time, thus making it difficult to estimate the amount of time spent in VR [27].

In this paper, we consider the disconnect caused by HMDs across the three aforementioned aspects of our lives – our physical awareness, our digital connectivity, and our perception of time – and explore notifications for physical, digital, and temporal events and/or interruptions. Although the term "notification" is commonly associated with digital events such as the arrival of an email or message, we use it as an overarching term to refer to cues (visual, auditory, or haptic) presented to immersed users to bring their attention to events occurring outside the virtual environment. Following the ephemeral nature of many digital notifications (e.g., a phone call only rings for 15-20 sec), we designed all notifications in this work to be short-lived and fleeting (i.e., 20 sec).

Much research in the VR field is aimed at bridging the gap between the immediate physical surroundings and the immediate virtual surroundings. While some researchers propose dynamic generation of virtual worlds based on real world objects [6, 28, 29], others suggest ways of folding the simulated world onto smaller real world geometry [32, 33]. However, the question of how to represent external events in the form of notifications in VR still remains largely unexplored. Popular notification methods, such as toasts, tickers, or icons [39] may be unsuitable or ineffective for a three-dimensional world. Voit et al. [38] suggested on-object and ambient notifications for smart homes. Similarly, an immersive virtual world provides a different set of affordances and constraints when compared to 2D interfaces.

• *Sarthak Ghosh, Lauren Winston, Nishant Panchal, Philippe Kimura-Thollander, Jeff Hotnog, Douglas Cheong, Gabriel Reyes, and Gregory D. Abowd are with the School of Interactive Computing at Georgia Institute of Technology. Emails: {sarthak.ghosh, lwinston, nish, philkt, jhotnog3, dcheong, greyes, abowd}@gatech.edu.*

Manuscript received 11 Sept. 2017; accepted 8 Jan. 2018.
Date of publication 19 Jan. 2018; date of current version 18 Mar. 2018.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TVCG.2018.2793698

In this paper, we aim to not only understand common interruptions faced by VR users, but also design notifications for them. We invited experienced VR industry professionals to generate a set of design possibilities. Fig. 1 shows examples of multi-modal notifications for a common interruption scenario. Further, we developed a Unity application to prototype some popular physical, digital, and temporal notification ideas. Each of the prototyped designs had visual, audio, and/or haptic components. To evaluate these designs, we conducted an empirical study on the noticeability and perception of 5 different VR notification designs across 6 modality combinations - audio only (A), haptic only (H), visual only (V), audio + haptic (AH), audio + visual (AV), and haptic + visual (HV). To conclude, we present a set of promising recommendations for VR notification design and development.

Our research aims to create a general understanding of common interruption scenarios faced by VR users. It provides VR designers with a set of possibilities and recommendations to consider while designing notifications. Further, our findings contribute to the body of knowledge on the performance of multi-modal notifications in virtual reality.

2 Related Work

In this section, we revisit literature on interruptibility, multi-modal notifications, ambient systems, and mixed reality. Further, we analyze existing commercial VR notification systems.

2.1 Multi-Modal Notifications and Feedback

Exploiting multi-modal output channels for notifications and feedback can enable new user experiences that capture users' attention, facilitate comprehension, and differentiate various levels of information. Past work in mobile and wearable computing have explored the use of multi-modal channels to notify users. NotiRing [26] uses multiple channels with a finger-worn ring to notify the user. They compare the performance of different channels such as light, sound, vibration, poke, and temperature under 3 different user contexts. Their results show that vibration has the quickest reaction time and the second lowest error rate, after sound. In an immersive VR context, vibration may be a useful channel for notifying users about an event without breaking immersion. Vibration may also be easily confused with diegetic tactile feedback. Song et al. [30] used multiple vibration patterns, squeezing, and temperature changes to generate different notifications with a wrist-worn wearable. Their results indicate that vibration was the least error prone, but unique timing patterns were indistinguishable. It is important to note that vibration alone may be insufficient to convey enough information, but may be sufficient to gain the user's attention.

Sonification is effective in conveying information to visually impaired users [9], as well as visually engaged users in automotive [14] and workplace settings [22]. Auditory displays can also be used for notifications, through the use of attensons (i.e., attention-getting sounds), directing the user to intermittent information [1]. Häkkilä et al. [12] suggest manipulations in audio variables such as length, vibrato, etc. to affect the perceived importance of the underlying event. Such manipulations should be considered and evaluated while designing VR audio notifications. In our studies, we have concentrated mostly on comparisons across modalities and their combinations. Mäkelä et al. [22] used walking sounds as auditory displays to indicate the presence of coworkers in an office environment. Haptic cues have been popular display choices for conveying navigational information, as well as warnings inside vehicles [3, 10, 35]. Nukarinen et al. [23] compared directional haptic and audio cues against visual cues in a vehicle navigation task. Their results indicate that non-visual cues achieve faster reactions while driving, although haptics result in more errors. Such results may not be directly extendable to a VR scenario because of varying design constraints and differences in visual, auditory, and tactile engagement.

Cidota et al. [7] explored how automatic audio and visual notifications about a remote user's activities affect collaboration in a virtual environment. The results show that visual notifications were preferred over audio or no notifications, independent of the difficulty level of the task. NotifEye [15] and eye-q [8] both utilize purely visual cues presented through eye-glasses to notify users about events in their digital ecosystems. However, these results may also not readily extend to more isolated and immersive situations in VR.

2.2 Interruptibility

Interruptibility has been studied extensively in the field of ubiquitous computing (e.g., driving, office, smart-home, etc.) [19, 34, 37]. In general, people can quickly assess how interruptible another person is in different situations. However, automating this process can be quite challenging. Most systems require the user to set preferences on their interruptibility – no notifications, priority notifications, or all notifications. The use of sensors to understand context-based interruptibility has been a popular topic for HCI researchers [31]. Hudson et al. [11] first demonstrated the idea of using sensor-based estimators to predict a person's interruptibility. However, their findings are limited by demographics and the office context which presents different challenges than the VR medium. Kern et al. [13] proposed 5 factors to define a user's interruptibility: the *importance* of the event that is being notified, the *activity* of the user, the *social activity*, the *social situation* and finally the *location*. In a VR context, many of these factors are less important since the majority of VR interactions are strictly personal and unnoticeable by others. The three factors that we consider to be relevant are the importance of the event, the nature of the task in VR, and the location - both physical (i.e., is the user at home, in the office or somewhere else?) and virtual (i.e., what does the virtual environment portray?). McFarlane et al. tested four known methods for interrupting people in an opportune way [19] – immediate, indirect, mediated, and scheduled. We can extend this work to the VR context. For example, we can imagine that the immediate solution is to directly reveal notifications using the HMD or other output channels. An indirect solution would be to indicate that a notification is ready and then allow users to decide when to access it. On the contrary, a mediated solution may embody an in-VR character (e.g., a virtual pet) or object to deliver the notification. Finally, we can deliver scheduled notifications at regular intervals to the user. In this work, we focus on encoding information across multiple synchronous channels with varied notification designs following the immediate, mediated, and scheduled mechanisms.

2.3 Ambient Awareness

Ambient or peripheral displays often serve as effective methods of maintaining continuous awareness about events that are not primarily in the focus of attention. Status indicators (e.g. battery life, connectivity, etc.) provide continuous ambient information about our devices. Many researchers have tried to move information off the screen onto physical objects, thus generating ubiquitous, aesthetic representations of data. Matthews [17] argues that notification displays differ from ambient displays as they are meant to attract immediate attention from users. However, according to other work [2, 24], some ambient displays can easily move from the periphery of attention to the focus, thus behaving more as a notification display. Drawing analogies to the real world, ambient displays can also be included in a virtual world to display both diegetic information (e.g., changing health of character, virtual time passing, etc.) and also to expose external pieces of information. Researchers have suggested dimensions for designing and evaluating different kinds of ambient displays [2, 24, 36]. In most cases, modality forms an important aspect of design and prior work agrees that an ambient display should use a modality that is not already overloaded. In a VR context, the visual, audio, and haptic channels are all often deeply engaged, with the visual channel being the most dominant. In this paper, we take a deeper look at the performance of different modalities and their combinations for notifications in VR. Our work is inspired by the model proposed by McCrickard et al. that encompasses interruption, reaction, and comprehension [18]. We measure reaction times and subjective comprehension for notifications in a VR experience.

2.4 Bridging the Physical-Digital Gap

Many researchers have focused on redesigning the VR world to suit the physical world. The lack of open tracked spaces lead to solutions like redirected walking [25]. Suma et al. [32] attempted to maximize walking in a small physical space by modifying the virtual world "behind the back" of the immersed user. Such approaches are suitable for controlled physical settings. Simeone et al. [29] proposed "Substitutional Reality" in which every physical object was paired with a virtual counterpart. Similarly, a depth map of the surroundings was used to dynamically generate and populate a 3D virtual space in "Reality Skin" [28]. Although these works provide excellent solutions for many environments, they do not account for situations in which the physical and virtual spaces are strikingly different in character, e.g., simulating an open virtual beach in a cluttered living room. In a slightly different approach, Haptic Retargeting [4] manipulated the perception of the virtual world by leveraging the dominance of the visual system when our senses conflict. This allowed the reutilization of the same physical object for different virtual counterparts. While these examples use information about or manipulate the physical world to enhance the virtual world, the external world is largely isolated from the VR scenes.

Exposing the physical within the virtual world can help bridge the gap between the two. Budhiraja et al.'s [5] enabled peripheral real world interactions by incorporating renderings of the physical world using transparency and overlays. Edging toward a broader interpretation of the term Mixed Reality [20], we also include in our explorations external sources of information, details about the physical surrounding (e.g., people speaking to you), the perception of time (e.g., time spent in VR), and a person's digital ecosystem (e.g., incoming phone calls).

2.5 VR Notifications

Recently, digital notification support has been added to the commercially available HTC Vive headset. Users must pair their smartphones with their computers over Bluetooth using the VIVE application (both Android and iOS). Users can then receive notifications about phone calls, text messages, and calendar events. Information is represented in the same way for all notifications - a blue panel is displayed at a fixed distance from the user, accompanied by a notification sound. Only the sender's name is provided and users must pause their VR experience to visit the VIVE dashboard for additional details. This non-customizable representation might not be suitable for all VR environments, thus degrading the experience. Zephyr[1] uses the local network to mirror all phone notifications to the VIVE dashboard. Although this solves various technical issues found in the VIVE notification system, the concerns related to the fixed representation of the notifications remain the same. The dashboard style interaction is also found in Google's Daydream ecosystem. According to recent announcements, all system generated notifications will be accessible in Daydream VR. However, the notification style is highly inspired from traditional 2D displays and might not be effective for a 3D environment. The same holds true for the GearVR and PlayStation VR experiences. Such developments provide encouraging evidence that the problem addressed in this paper exists and needs to be further explored.

3 Common Interruption Scenarios

We conducted a Qualtrics-based online survey to understand the common interruption scenarios that people face and how they feel about mapping these scenarios into their VR worlds. Through online VR forums and university mailing lists, we reached out to a broad population. We received 90 responses and out of them 61 (only those who had used immersive VR equipment) were considered valid.

3.1 Survey Design

The survey begins with a section of demographics questions. In this section, we also asked participants the maximum amount of time they have spent in VR at a stretch. Section 2 included three questions similar to the following question, to understand more about the users' needs for physical, digital, and temporal awareness – "Can you tell us about some incidents when you wanted to be aware of your "real" physical surroundings while you were immersed in a Virtual Environment? You can choose all that applies from the examples given and also provide other incidents." Each question provided 4 examples and left space for two open responses. Some examples provided were: "I stepped on something on the floor," "I heard a knock on the door," "I was approaching a wall," etc. In Section 3, we asked people to rate a set of interruption scenarios based on their relevancy – "How relevant do you think it is to notify you about the following situations that may arise while you are immersed playing in VR?" The question provided 16 scenarios each of which could be rated on a scale of 1-5, with 1 being "Extremely Irrelevant" and 5 being "Extremely Relevant."

3.2 Survey Results

Demographics. Participants belonged to a wide range of professions including students, researchers, architects, VR developers, designers, and consultants. They represented a broad set of age groups – 18-25 (23.1%), 26-35 (48.1%), 36-45 (19.2%), 46-55 (7.6%), and 56-65 years (2.0%). On a 5-point Likert scale, all accepted participants deemed themselves to be either completely familiar (76.9%) or somewhat familiar (23.2%) with VR systems. Interestingly, while most people (33.0%) reported having spent less than 20 minutes in VR at a stretch, almost 22.0% people reported having spent more than 2 hours. This hints at how quickly the technology is progressing.

Questionnaire. People were more concerned about their immediate physical surroundings than about missing digital notifications overall. We hypothesize this is understandable because commercial room-scale VR is still in an early stage and our homes are not specifically designed for VR experiences. People often face problems like "bumping into a chair" or "walking into a wall." 36 people (59.0%) mentioned they wanted to be better aware of walls; 32 people (52.4%) wanted to be aware of what they were stepping onto. Other than that, the changing physical environment also seemed a major concern for 31 people (50.8%), including cats coming into the play area, and people entering/leaving the area without the player realizing. On the digital front, many people reported the need to check phone calls (39.3%) and text messages (31.1%). Interestingly, 38 people (62.3%) mentioned that they have experienced lack of time perception in VR.

We also received several interesting comments in the open-ended responses in Section 2. Many people complained about the interruptions caused by the cord attached to the HMD, which may disappear as wireless headsets become readily available. While several comments explained their choices in the survey, others revealed new discomforts or interruptions. For example:

> "I was standing near a couch. When I turned to look at something behind me I bumped into the couch."

> "I felt weird being immersed in VR when I knew other people were in the physical room."

> "A loud noise in the real world should cause the automatic pause or decrease the volume in my virtual env[ironment]."

> "My phone was ringing and I missed the call. It was quite an important call. I would like an app that shows who is calling on the phone on my hmd [head-mounted display]."

Notification Scenarios	Mode Relevancy	Count
You are very close to a wall	5	42 (68.85%)
An emergency alarm rings	5	42 (68.85%)
Someone in your room is calling your name	5	33 (54.10%)
Someone is about to tap your shoulders	5	26 (42.62%)
You step on something on the floor	5	22 (36.07%)
Some one is giving you instructions	5	22 (36.07%)
Someone enters your room	4	31 (50.82%)
Some one is knocking at your door	4	29 (47.54%)
Your phone alarm is ringing	4	28 (45.90%)
You are receiving a phone call	4	26 (42.62%)
You have an appointment in 15 minutes.	4	24 (39.34%)
You have received a text message	4	21 (34.43%)
Your door bell just rang	4	21 (34.43%)
You want to know how much time you have spent already	3	23 (37.70%)
You have received an email	1	16 (26.23 %)
You have a new Facebook notification	1	45 (73.77%)

Fig. 2. Likert scale relevancy scores (modes) of different scenarios based on a survey with 61 VR users. Count : number of people who gave the mode score.

[1] http://gaubert.io/zephyr/

In Fig. 2, we can see how each scenario was rated on a relevancy scale. The first two scenarios are intuitively understandable because people naturally want to avoid accidents. On the digital front, phone alarms, incoming calls, and calendar notifications received relatively higher ratings than social media notifications. We synthesized the results of the survey and the comments to create a list of 10 notification scenarios that were not only highly relevant but also covered a wide range of situations. See Fig. 3 for the full notification list. The notifications include 4 relevant and popular scenarios belonging to each of the physical and the digital categories, and 2 scenarios belonging to the temporal category. We assign each scenario an abbreviation code to be used as a reference in the remainder of the paper.

4 Designing VR Notifications: A Design Exercise

We conducted a co-design exercise with professional VR content designers and developers to explore the possibilities of VR notifications. One of the common challenges in co-design is participants' limited knowledge of the underlying technology. Research methods that are "participatory" in nature need to ensure that participants have prior exposure to the "target implementation technology" [21], which we considered in our choice of participants. We followed an "elicitation" study procedure to gather a wide variety of notification designs, and also understand preferences and associated concerns. We conducted a pilot within our research group to finalize the appropriate procedure.

4.1 Participants

The design exercise was conducted with 7 participants (6M and 1F). All participants were experienced VR content developers and/or designers in our institution's metro area. Their professions included VR game development, VR film-making, VR marketing, and consulting. 6 of the 7 participants owned personal VR setups at home.

4.2 Procedure

The entire design exercise was conducted in 2 hours and was video-recorded. Introduction and task explanation lasted 30 minutes, the design session was 60 minutes, and discussion was 30 minutes. Sticky notes and sketches were collected as outcome artifacts.

4.2.1 Introduction

All participants were seated together and given a short introduction to the project. They were told to consider three output channels (e.g., visual, audio, and haptic), while thinking of VR notifications.

4.2.2 Designing

The pilot exercise had revealed that notification representations may vary depending on the virtual environment that they are designed for - e.g., a virtual living room may suggest different design possibilities when compared to a virtual desert environment. Other properties of an environment, such as clutter, interactivity, lighting, etc. may also influence notification design. To restrict our project's scope, we decided to consider two different types of VR environments (*E1* and *E2*, see Fig. 4) for the design session. *E1* represented a closed and busy environment with surfaces and objects around the user (similar to Job Simulator[2]). *E2* represented a fairly open environment with undefined boundaries (similar to Tilt Brush[3]). *E1* and *E2* are representative of a wide range of popularly available VR experiences.

Scenario	Category	Code
People entering your physical room, when you are in VR	Physical	PPM
Someone is trying to get your attention from your left	Physical	PAG
You are about to bump into a wall towards your right	Physical	BWR
You are about to trip over your pet	Physical	PET
You have an incoming call	Digital	PHO
You have an email from your boss	Digital	BOS
You have a text message from your friend	Digital	FRI
You have a text message from your spouse	Digital	SPO
You have already spent an hour in VR	Temporal	HOU
You have an appointment approaching	Temporal	APO

Fig. 3. List of 10 relevant & encompassing VR notification scenarios

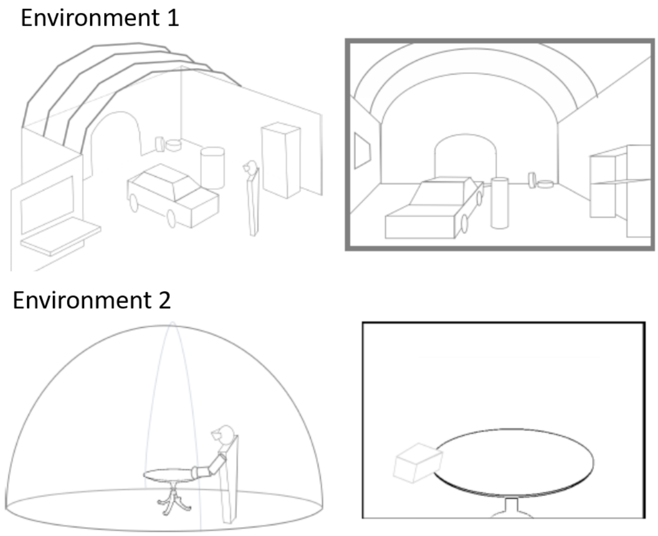

Fig. 4. Sketches as thinking props. Top: E1. Left shows a person in a closed garage-like space. Right shows POV representation of the scene. Bottom: E2. Left shows a person in an open environment - the user is making a virtual table. Right shows the POV representation of E2.

Participants received the list of 10 interruption scenarios to design for, shown in Fig. 3. The list was compiled based on the survey results from Section 3. Each participant was asked to come up with notification designs for all scenarios – 5 for *E1* and 5 for *E2*. The order of scenarios were randomized across participants to ensure each participant could design for both environments. Following Weber et al. [39], sketches (see Fig. 5) were used as thinking props. Each sketch showed an environment (*E1* or *E2*), a VR headset, controllers, and a pair of hands. Participants could write/draw on the sketches. They were asked to write their design rationales on separate post-it notes, to be analyzed later.

4.2.3 Discussion

After the design session, all participants placed their post-it notes by scenario on a wall visible to everyone. While placing each note, the participant explained their design rationale to others. Often times, this generated follow-up discussions, concerns, and affirmations.

4.3 Results

All post-it notes were rearranged based on audio, visual, and/or haptic modalities and subcategories were identified. The recorded video was also transcribed. Analyses of the transcript, paired with the categorization of post-its, helped us identify several emergent themes.

Overall, participants engaged in a lot of discussion around the theme of content display. Representing "enough" information in a notification is always a challenge. While some argued that revealing as much information as possible is desirable for physical notifications, others felt that the aesthetics of the VR environment must be considered. In a relatively undefined environment (e.g., Tilt Brush) detailed visuals can be utilized for physical notifications without hampering the user experience. The same needs to be controlled for a more defined and confined environment (e.g., Job Simulator). For digital notifications, the amount of information to be displayed depends on the importance of the underlying event and should be customizable.

In the following subsections, we present different design concepts categorized by their modalities. Words and phrases spoken or written by our participants are mentioned in quotes.

[2]http://jobsimulatorgame.com/
[3]http://www.tiltbrush.com

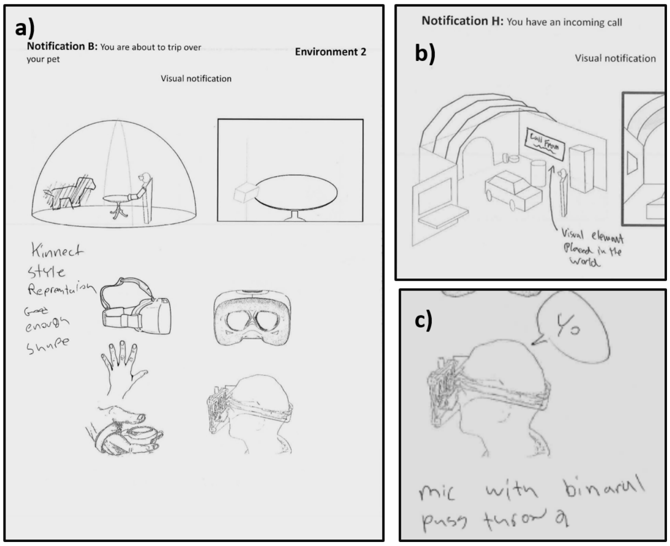

Fig. 5. Sketches done by participants. a) Participant sketched a dog in the environment and wrote "Kinect style representation" for a physical interruption. b) Participant drew a banner on the wall and scribbled "Visual element placed in world." c) Participant sketched a bubble around the headphones and scribbled "Mic with binaural pass through."

4.3.1 Audio

Spatialized Sound. 3D localized sound is at the heart of VR content and should extend to notification design. For example, when a person is approaching the immersed user, "footstep" sounds can be mapped to the exact or approximate location of the moving person. The idea can extend to reveal the position of a pet, possibly replacing footsteps with other identifiable sounds. Pre-recorded voices of friends or family members and even pet sounds can be spatially mapped to quickly convey information (in the background) about the presence of another living being in the play area. Similarly, identifiable voices can be used to announce or read aloud notifications like text messages. Another technique is to introduce audio from the real surroundings (line in) whenever another person or a pet enters the play area.

Skeumorphism. Many designers preferred metaphorical designs for notifications. People entering and leaving the play area can be represented through creaking sounds of a door. Just like the door metaphor, the clock metaphor was a common suggestion for temporal notifications - "utilize grandfather clock hourly chimes" or "utilize the alarm-sounds of your watch." With digital notifications, it makes sense to re-utilize existing notification styles from the phone. For example, we can mimic the phone and play the corresponding sounds for text messages, incoming calls, emails, etc. Users should have the capability of customizing their tones, based on what they are used to in real life.

Gradual Intensification. Gradually intensifying sound effects can be used to provide a sense of danger approaching. Suggestions included "start slow and intensify later, eventually freezing the game." This seemed specifically suitable for accidental situations, such as approaching a wall or tripping on a pet.

4.3.2 Haptic

Skeumorphism. Feet vibrations were suggested as a means to inform users about pets and children entering the play area. While a suggestion that vibrations can be matched to the footsteps of an approaching person was shared, different persistent vibration patterns could also be used to indicate that "you need to freeze" in your position. Furthermore, tap-like sensations on the shoulder could be great for certain physical interruptions to grab your attention. Similarly, vibrations on the ring finger may indicate messages from the person's spouse. Controller vibrations could be utilized to provide digital notifications for phone calls or emails. Similar to visual and audio notifications, participants suggested the same vibration patterns commonly used in our phones should be extended to VR notifications.

Directionality. Vibration motors could be attached to different parts of the headset to provide a sense of directionality (e.g., "look left, or look back"). For highly critical notifications, all vibrations of the head and controller could be triggered simultaneously. Moreover, left and right directionality can also be provided by vibrating the left or right controllers, respectively. Such vibrations can serve either as a "pre-notifier" or to direct the user's attention somewhere in the scene.

Gradual Intensification. Gradually intensifying sensations can be used to indicate "approaching danger." Parallels were drawn to the "spidey sense" superpower in the comic hero Spider-Man. One participant mentioned "slight buzz as you approach and becomes more intense as you come near," an "...ultrasonic wave like system to give you the feeling of a wall even before you reach it."

4.3.3 Visual

3D Representation. The most natural choice to notify a user about extraneous physical activities in the play-area was through different 3D representations. For people and pets entering the play area, popular suggestions included representing their 3D models (generated with real depth data) in VR. When somebody tries to talk to the immersed user, a "hand wave" can be used to grab their attention. While many participants wanted to use detailed models similar to Intel's new "Merged Reality" technology[4], others went for a more minimalist approach. "We can use grids to show the presence of others," or "use a glowing orb as placeholders for people within a 3 feet radius from you." "Slow fading in" animations can be used to prevent users startling because of the sudden appearance of avatars. Animations also play a part in creating a distinction between the real and the virtual world. "You can use echo lines to indicate their presence (like ripples of light in concentric circles)." For static objects, the currently available "chaperone system" (i.e., a translucent grid to indicate play-area boundaries available in the HTC Vive) is a preferred choice. For example, "we can piggyback off the chaperone system and show the presence of furniture and walls (only when you are near them)."

Skeumorphism. If the environment allows, a clock model can be placed on a suitable wall to show the real world time. In other cases, a "smartwatch" can be placed on the user's virtual hands. One participant suggested the smartwatch should only be activated when the user wants (e.g., by tapping their wrists) and be hidden otherwise. For digital notifications, the controller on the non-dominant hand can be used to provide familiar phone-like interfaces. For example, "present a ringing phone with caller information on a controller."

3D Pop-Ups. Pop-ups or small billboards were common choices for "digital notifications." For open environments, "3D pop-ups floating towards the back of the world or on existing surfaces" make sense. However, for busier environments, pop-ups near the controller were popular choices. Some examples included "display like a side bar beside controller," "shaded notification on a controller," and "pop-up in the same way as a phone does." Another suggestion for critical notifications was to pin a pop-up to the camera's field of view.

[4]https://newsroom.intel.com/chip-shots/intel-unveils-project-alloy/

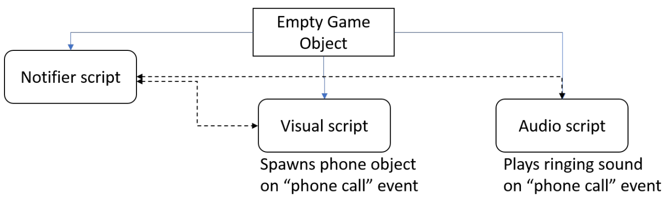

Fig. 6. Notifier, visual, and audio scripts are attached to an empty game object. The Notifier script turns the empty object into a Notifier Game Object. The Visual and Audio scripts both subscribe to the "phone call" event to define visual and auditory behaviors, respectively.

Flash and Pause. Ambient notifications were suggested for "collision" notifications. For example, "flash your environment" when you are about to bump into a wall. Peripheral lights were also suggested for providing directional information, such as "have a small notification on the side of the view so that you know which direction to look at." For critical situations (e.g., on the verge of tripping over your pet), pausing the game and revealing the real world seemed like the way to go, basically "freeze the place." Parallels were drawn to the way the Google Chrome browser handles calendar notifications, shifting your focus to a new tab and forcing you to interact with the notification.

5 NotifiVR Unity Application

The design exercise resulted in a variety of notification design concepts, which could vary largely in their content, look, and feel based on the VR scene under consideration. Hence, we made the logical choice to focus on different modality combinations that may arise across variations in designs. Based on insights gained from the exercise, we went through multiple sprints of sketching and ideation to identify a varied set of visual, haptic, and audio notifications to prototype. These notification concepts were directly related to the ideas suggested by experienced VR professionals. In order to create functional prototypes, we developed a Unity-based application that integrates audio, visual, and haptic notifications into a scene.

The NotifiVR application expects network events from the external world as inputs. We defined a Notification Manager prefab (i.e., a type of game object in Unity) that listens for triggers from networked sensors. The Notification Manager also lists all the Notifier Game Objects in the scene. A Notifier Game Object is a game object that has a Notifier script attached to it. Such an object can then have multiple Visual, Audio, and Haptic output scripts. Each output script can be used to subscribe to a particular notification event by using drop-down lists. Once subscribed, the desired behavior for the event can be customized. For example, in the event of a phone call (see Figure 6), visual and audio output scripts can both be attached to the same notifier object. Upon receiving the event, the visual output script may spawn a new phone object and the audio output script may play a ringing sound.

In its current version, the NotifiVR application can successfully receive digital events and information from a companion Android application over TCP. We understand the need for substantial tracking of the user's space to be able to sense most physical interruptions. In this paper, we use a Wizard of Oz approach to feed simulated event information from the physical world into our notification manager.

5.1 Notification Designs

We implemented 5 different types of notifications belonging to the physical (2 designs), digital (2 designs), and temporal (1 design) categories. Each design had a visual, audio, and haptic component. All audio notifications used localized sounds and all haptic vibrations were provided through the controllers. The details of the designs are provided in Fig. 7. For the physical category, the scenarios chosen were "a person entering the play area" (PPM) and "a person trying to talk to you" (PAG). Other scenarios involving pets or walls easily extend from these situations. Both notifications would appear within a 3 meter radius of the user, to simulate real people in the play area. We prototyped two digital notifications for an incoming phone call (PHO) and an incoming Slack message (POP). The visuals for the POP notification would appear on existing surfaces near the users, taking into account their momentary position and direction. We restrained from directly attaching any notification to the user's field of view as we considered it to be too intrusive and obviously noticeable. For the temporal scenario, we prototyped a wrist watch (WAT) to notify participants of the amount of time spent in VR.

The designs were chosen to represent a wide variety of possibilities on the body, on the controllers, and in the environment. The prototypes were kept simple and easy to comprehend, maintaining familiarity with existing real world concepts. In the future, it is possible to also appropriate other real-world items for VR notifications (especially for haptics), such as smartwatches and phones in pockets. The aim of our empirical study (see next section) was to primarily evaluate the role of modalities in different VR notifications, not to compare the entire space of all design possibilities.

6 Empirical Study

We conducted a user study using 5 designs of VR notifications presented across 6 different modality conditions. Through the user study, we hoped to understand how the use of different modalities can affect the reaction time of notifications. We also wanted to evaluate how people perceive 3-dimensional notifications in general.

6.1 Participants

We recruited 16 volunteer participants (8F and 8M), ranging from 20-47 years (M=26.6, SD=6.2). Six participants were interaction designers. 4 people (25.0%) reported having used VR systems for 30 minutes to an hour at a time, while one person reported having used it for 2 hours straight. Other participants (68.75%) had used it for less than 30 minutes at a time. Most participants were not frequent users of VR and the average score of VR familiarity was 1.81 (1 being completely unfamiliar and 5 being completely familiar). Only one participant was highly experienced with VR technologies.

6.2 Virtual Environment

We used freely available assets to create an "abandoned warehouse" virtual environment for the user study. An exhaust fan was used as a source of ambient sound. To break the monotony, we introduced intermittent sparking effects. Visual effects of floating dust particles were also added. To understand the average level of presence supported by the environment, we used the revised "Presence Questionnaire" by Witmer et al. [40]. The presence score on average was rated 137.8 (SD = 15.8) by 16 participants, out of a possible score of 172. The breakdown of the presence score includes realism (M = 39.7, SD = 5.5), possibility to act (M = 25, SD = 2.8), quality of interface (M = 17.4, SD = 2.5), possibility to examine (M = 16.6, SD = 3.0), self evaluation of performance (M = 12.6, SD = 1.3), sounds (M = 16.8, SD = 3.6), and haptics (M = 9.8, SD = 3.2). These scores indicate that the complete experience in the environment provided a considerable level of presence. In this paper, we utilize the scores as guidance to quantify participant experience with our environment and to inform future studies.

6.3 Tasks

We evaluated the performance of modalities in notifications within the context of a primary task. The task was designed to include common VR interactions such as visual search, exploration, navigation, grabbing, and throwing. Our task required users to explore the environment and find a ball floating 1m above the floor. The spawn of the ball was accompanied by a localized audio cue and a white arrow on top. This sound effect ensured that users do not rely completely on visual cues while searching for the newly spawned ball. Moreover, in-game audio cues are quite common for VR experiences. Users had to navigate to the ball by "teleporting," using the left-trigger button. Once near the ball, users had to grab the ball with the right-trigger button and look for the target (i.e., a glowing yellow ring floating 1m above the floor). Users were required to navigate to the target, while holding on to the ball. The task ended with the user throwing the ball into the target, for which spatialized audio feedback was provided. Grabbing and releasing the ball were accompanied by mild vibrations of the right controller. To create a distinction, the intensity of vibrations was intentionally kept lower than that of the external haptic notifications. Each task was available for 30 seconds. In the case of non-completion, the ball and the target would simply disappear and a new pair would appear at random locations. The primary task remained the same across participants for consistency and to ensure the duration of the study remained within acceptable limits below two hours. See the attached video figure for a demonstration of the environment and tasks performed.

Following the example of NotiRing [26], we asked participants to "acknowledge the notification" as their secondary task. Users could acknowledge a notification by pressing either of the touchpad buttons (left or right). Each notification was presented for 20 seconds, following the

Person talking (PAG)
Audio: Human voice
Haptic: Closest controller; "knock knock" pattern
(a)

Incoming call (PHO)
Audio: Phone Ring.
Haptic: Left vibrations mimicking the audio
(b)

Slack pop up (POP)
Audio: Slack notification.
Haptic : Closest controller; mimicking the audio
(c)

Person in room (PPM)
Audio: Footsteps on carpet
Haptic: Alternative left-right controllers
(d)

Time spent (WAT)
Audio: Clock strike.
Haptic: Left vibrations mimicking the audio
(e)

Fig. 7. Visual, audio, and haptic components of all the notifications prototyped – a total of 15 designs (i.e., 5 events with 3 output channels each).

example of a standard ringtone. During this time, visual notifications were continuously displayed in the environment, while audio and haptic notifications were toggled periodically (see Fig. 7). The vibration intensity was constant but the patterns differed for each notification. Similarly, the gain and attenuation for each audio source were kept constant. Pilot studies suggested that all audio sources were easily audible, and the vibration intensity felt comfortable and non-jarring.

6.4 Design

The experiment followed a 5x6 within-subjects design. The two independent variables were the notification type (e.g., person talking in room (PAG), phone call (PHO), Slack pop up (POP), person in the play area (PPM), and time spent (WAT)) and the notification modalities (e.g., visual (V), haptic (H), audio (A), visual + haptic (HV), audio + visual (AV), and audio + haptic (AH)).

The measured dependent variables were reaction time, task completion time, reaction error rate, and task error rate. Reaction time is the time to acknowledge a notification from the moment it became active. Task completion time is the time to complete a task since a ball became active. Reaction error rate is the number of times participants didn't respond to a certain notification, out of all notification activations. Task error rate is the number of times a task wasn't completed, provided a certain notification was shown during the task. Task error rate was later found to be insignificant and is not discussed further in the paper.

6.5 Procedure

Following an introduction, each participant answered a demographics questionnaire. They were given a tutorial of the HTC Vive and its controllers. Participants were then immersed into our environment by putting on the headset and over-the-ear headphones.

Training period: The training period consisted of three parts. In part 1, participants were asked to teleport three times. In part 2, participants completed the ball task 3 times consecutively. In part 3, they experienced all notifications one after the other (30 conditions) and were asked to express what they saw, heard, or felt.

Testing period: There were 2 test sessions with a 5-minute break in between. Each session consisted of 45 ball tasks. Out of the 45 tasks, 30 were randomly interrupted with notifications. Notifications were presented in counterbalanced order, according to a Latin square of 30 conditions (5 notifications x 6 modalities). Thus, each participant had two chances (once in each session) to respond to all 30 notifications presented in two different orders.

6.6 Post-Test Questionnaire

The post-test questionnaire contained four 7-point Likert scale questions for each of the 30 notifications: *Noticeability* (i.e., how easy or difficult is it to notice the notification?), *Understandability* (i.e., once you notice the notification, how easy or difficult is it to understand what it stands for?), *Perceived Urgency* (i.e., what level of urgency does the notification convey?), and *Perceived Intrusiveness* (i.e., how much of a hindrance was the notification to the overall VR experience?).

We collected user preferences through a notification matching task. Participants were presented with the same scenarios that were used for the design exercise reported earlier in Section 4, and were asked to indicate their choice of modality for each. They were also asked a few open-ended questions to understand the rationale behind their choices.

6.7 Results

We analyzed reaction time, reaction error rate, and task time for all notifications. We also analyzed the post-test questionnaire data to gauge how participants perceived different notifications and their modalities.

6.7.1 Reaction Time

A two way repeated measures ANOVA test across modalities ($F_{5,75} = 37.67$, $p < 0.01$, $\eta^2 = 0.35$), followed by a paired t-test with Bonferroni correction reveals significant differences between visual (V) and all other modality conditions (all $p < 0.01$). The headset's limited field of view made it difficult to spot purely visual notifications. Participants stopped looking at their hands after a while and visual notifications on the controllers were easy to overlook. Consistently, the highest average reaction time is for the V modality for all notifications (see Fig. 8). Although the differences are not significant, we see that AH and HV, consistently achieved the quickest reactions. Audio and haptic are both effective in getting a reaction out of participants in VR. This being said, we should note that combinatorial modalities involving audio and/or haptics (AH, HV, AV) have most often performed better than the pure modalities of visual, audio, and haptic. For AV and A, we see that the person in the room (PPM) notification took the longest to respond to. This is expected because depending on where the footsteps were (between 1-3m from the user), they would be relatively faint or loud.

We also found a significant effect on notification type ($F_{4,60} = 2.65, p < 0.05$, $\eta^2 = 0.02$). Figure 8 demonstrates that the phone call (PHO) consistently performed better than the person in the room (PPM) across all modalities. A pairwise t-test with Bonferroni correction reveals significant differences between PHO and PPM ($p < 0.01$). Reaction error rates are significantly high for all the visual (V) conditions. For the temporal notification (WAT), we see non-zero error rates for AH and HV as well. On-wrist notifications may not be the best choice for active notifications, but can be useful to provide on-demand non-intrusive information the user can check when needed. Although on-hand notifications were often difficult to spot, they were also deemed to be more credible and easy to associate with external events. Many users appreciated the use of the controllers for the incoming call notification (PHO). A participant also mentioned that since the controller is part of both realities, when a notification arrives on the controller, it becomes easier to associate it with the "real" reality.

We found no significant interaction effect between Modality and Notification types ($F_{20,300} = 0.88, p > 0.05, \eta^2 = 0.03$).

6.7.2 Task Time

No significant differences were found in the task completion time based on the notifications shown. From Figure 8 we observe that non-visual modalities like audio + haptic (AH), purely audio (A), and purely haptic (H), had slightly higher task times than others. Many participants did not attempt to search for any visual notifications which

Notification	Reaction Time (s)						Error Rate (%)						Task Time (s)					
	A	AH	AV	H	HV	V	A	AH	AV	H	HV	V	A	AH	AV	H	HV	V
PAG	3.08	3.01	3.17	4.66	**2.78**	7.23	**0**	**0**	**0**	12.5	**0**	3.13	**12.49**	15.99	13.94	14.61	14.2	14.1
PHO	3.01	**2.49**	2.65	3.91	3.11	6.42	**0**	**0**	**0**	**0**	**0**	40.63	15.14	13.96	14.72	14.8	**13.24**	14.81
POP	3.95	3.43	3.63	3.75	**2.69**	7.65	6.25	**0**	**0**	**0**	**0**	21.88	14.61	**13.45**	13.88	15.61	15.4	13.69
PPM	4.51	**3.31**	4.31	4.04	3.39	8.57	**0**	**0**	**0**	**0**	**0**	18.75	14.2	15.19	14.46	14.95	**14.15**	14.22
WAT	3.05	**2.78**	3.23	3.01	3.25	8.84	**0**	3.13	**0**	**0**	3.13	65.63	13.97	14.4	**13.35**	14.37	15.74	14.9
Overall	3.52	**3**	3.4	3.87	3.04	7.74	1.25	0.63	**0**	2.5	0.63	30	14.08	14.6	**14.07**	14.87	14.55	14.34

Fig. 8. Summary of reaction times and reaction error rates. Best values for each notification design are in bold and second best values are underlined.

led to interruption-free tasks. Other participants always performed a quick visual search to see if there were any visual representations of notifications. This short period of reorientation might have also contributed to higher task times.

6.7.3 Perception

We performed Friedman's test on Likert scale data (see Fig. 9) related to modality and notification type factors for each of the attributes: noticeability, understandability, perceived urgency, and perceived intrusiveness. We found significant effects of both the factors on each of the 4 attributes ($p < 0.01$). We further performed pairwise comparisons using Wilcoxon signed-rank tests with Bonferroni correction.

Noticeability. Pairwise tests on modalities revealed significant differences between purely visual (V) and every other condition (all $p < 0.05$), and between haptic + visual (HV) and audio + visual (AV) ($p = 0.016$). Purely visual notifications suffered from noticeability across the board. This effect can be partly due to the engagement in the task, and due to the lack of directional cues. Results also suggest that audio is a better companion to visuals than vibrations. Participants reported that because of pre-existing familiarity with non-diegetic (real world) audio notifications, these can be easily distinguished from diegetic audio cues. Haptic notifications were more difficult to distinguish from diegetic haptic feedback. We found significant differences between the Slack pop-up (POP) and person talking (PAG) ($p = 0.012$), POP and incoming call (PHO) ($p = 0.038$), PAG and time spent (WAT) ($p = 0.012$). These differences show how the design of certain notifications can increase their noticeability. The person talking (PAG) involved the appearance of a mesh representation of a person, which due to its large size paired with distinct "voice" was clearly noticeable, and often termed as "jarring." However, the watch (WAT) notification was difficult to notice because it often fell outside the field of view. The pop-up (POP) would appear on surfaces near the user. This made it very easy to overlook when the user was on the move, or teleporting. The vibrations alone failed to provide enough directionality and left users confused.

Understandability. Although vibrations were easy to feel, they were difficult to comprehend. Vibration patterns were not distinguishable. A participant mention "when you shoot or kill in a game, the controller usually vibrates. If there is continuous haptic feedback coming then it will be difficult for out of VR notifications." Participants preferred AH ($p = 0.01$), A ($p < 0.01$), AV ($p = 0.01$), HV ($p < 0.01$) and V ($p = 0.01$) in terms of understandability. In contrast, the purely haptic representation of the person in play-area (PPM) was rated high in understandability. All participants mentioned that "the left/right vibration" was relatable to the walking metaphor. The time spent (WAT) notification suffered from understandability. The metaphor of clock strikes wasn't suitable for quick understandability, as the participant had to wait to understand the time being represented by haptic or audio channels. We observed significant differences between WAT and PAG ($p = 0.012$), WAT and PHO ($p = 0.041$), WAT and PPM ($p = 0.011$).

Urgency. Conditions involving different combinations of haptic/visual/audio were perceived as more urgent in general. We found significant differences between H and AH ($p = 0.04$), H and AV ($p = 0.029$), and V and AV ($p = 0.04$). Interestingly, the audio channel

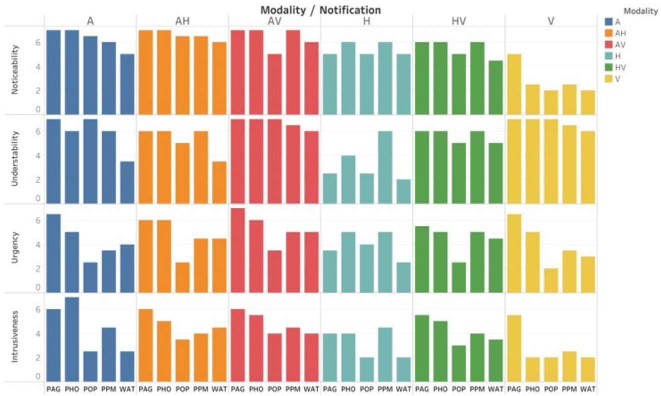

Fig. 9. Median Likert scores per modality per notification

in a multichannel notification seemed to add to the perception of urgency, even more than the haptic channel. Some participants associated audio with real life alarms and hence felt that it represented urgent situations. Across the notification type, the person talking notification (PAG) was perceived to be significantly more urgent ($p < 0.05$) than pop-ups (POP), person in play-area (PPM), and the time spent notification (WAT). The appearance of a life size mesh and the audio of a human voice seemed to demand immediate action.

Intrusiveness. Audio was again perceived as an intrusive component. Purely audio (A) notifications were rated as more intrusive ($p = 0.02$) than purely haptic notifications (H). Similarly, audio + visual (AV) is considered more intrusive than both H ($p = 0.01$) and V ($p < 0.01$). Among the notification types, person talking notification (PAG) was considered to be more intrusive than all other types (all $p < 0.05$). Moreover, because of the continuous head movement of the participants, they were often surprised by the appearance of the mesh (PAG). Incoming call (PHO) was significantly more intrusive than both POP and WAT. According to participants, this had to do with the closeness of the notification to their body, and to its relatively higher noticeability over WAT and POP. The perceptions of intrusiveness and urgency seem to be positively correlated ($r = 0.51$).

6.7.4 Matching Modalities to Scenarios

In the notification matching task of the questionnaire, participants were presented with the same scenarios that were mentioned earlier in Section 4. Participants were asked to indicate their choice of modality for each situation. Figure 10 represents the results of user preferences and shows how combinatorial modalities, such as AV, HV, AH and AHV (all three channels of audio, haptic, and visual) were largely preferred over the purely visual, audio, and haptic channels. As revealed by the open-ended questions, audio was preferred over haptics as an attractor of attention. Visual was preferred as a provider of finer grained information than what a audio notification could provide, for example, "to judge whether I should ignore it." Accordingly, we observe the high

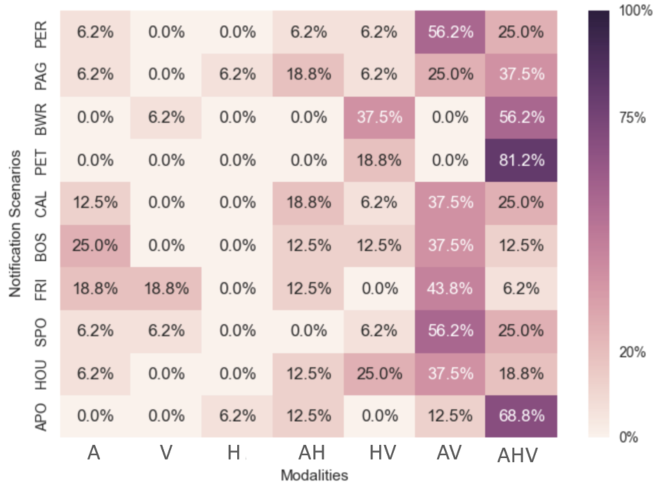

Fig. 10. User preference of channels per scenario

preference of AV in Figure 10.

Participants also based their choices on the perceived importance of a notification scenario. Pet entering (PET) and appointment approaching (APO) were considered highly important, requiring immediate attention; emails (BOS) were considered less urgent. Hence, we see people choosing AV or A over AHV. Users also tried to carry across their real-world experiences and familiarities. For example, "if a person is trying to talk to me, I would probably see and hear him first, so I chose AV" (P8). We notice AV to be a popular choice for both person entering/walking in the play area (PPM) and person talking (PAG). Curiously, for the bumping into a wall notification (BWR) and PET, participants showed relatively more support for HV (compared to AV), possibly because these situations in real life wouldn't involve much sound. AV was also the most popular choice for all digital scenarios, because "that's how it is usually in the phone" (P11). Some participants also went for pure audio when they felt the situations didn't demand immediate visual attention – emails (BOS) or text from friend (FRI). The only time visual (V) was somewhat preferred was for FRI, mainly due to the seemingly low consequences of missing such a notification.

7 Discussion and Design Recommendations

Our empirical study showed that purely visual notifications had poor noticeability most likely due to the highly visual nature of VR. While haptics were more difficult to understand, they were successful for on-controller notifications. Audio was most preferred for directing attention to the visuals of a notification. Combining these results and participant feedback with the outcomes of our design exercise, we arrived at the following design recommendations for VR notifications.

Creating Distinction. Notifications for external events must be distinguishable by design, as the VR environment can generate its own multi-modal diegetic feedback. VR notifications can be distinguishable by following a visual design vocabulary that is different from that of diegetic objects (e.g., meshes to represent physical obstacles). Similar guidelines can be found in related work on ambient displays [2, 24, 36]. There must always be an effort to make the ambient display standout from its environment. Matthews resonated a similar idea, stating "make visuals distinct" for glanceable displays [17].

Two participants mentioned that our notifications looked very different from the other objects in the warehouse, which made them recognizable as external notifications. For audio notifications, adding customizability to notification sounds is recommended. All participants mentioned that it was problematic to keep count of the number of times a sound/vibration was played to indicate time (WAT). Thus, information should not be coded into the repetition pattern of audio or haptic notifications. If the VR experience has diegetic haptic feedback, then creating a sharp distinction between the haptics used for feedback and notifications is challenging. It is advisable to either use haptic for a few unique notifications (e.g., emergency alarms) or use a completely separate haptic device for the notifications (e.g., a wrist band). One participant said "...I thought the vibrations should be somewhere else, other than the controllers. May be on the wrist, or on the back as a tap."

Using the Controllers Effectively. Notifications can be designed to appear anywhere in 3D space in a virtual environment. Thus, the 3D position of a notification with respect to one's body is also important to consider. The controllers serve as effective locations for notifications. However, one challenge is that controllers frequently fall outside the field of view. Thus, purely visual notifications on the controllers are easily missed. Although vibrations were rated low on understandability, they could be used as low-intrusive ways of drawing attention to the controllers for haptic + visual notifications. One participant said "I kept looking at the controller every time there was a vibration, instinctively." One participant observed that controllers were "both real and virtual," meaning one could see an exact virtual representation of the real controllers. Thus, it was more credible to receive real world notifications on them. However, since controllers are used for many actions in VR, they should be reserved only for high priority notifications. Further, controller vibrations are not good indicators of directionality.

Reducing Visual Search. Visual search for notifications in the environment can be reduced by using binaural audio to guide the gaze, or by using a special dedicated place/object for showing notifications. Spawning notifications in the middle of the user's field of view can reduce visual search, but can be highly intrusive and might require pausing the VR experience. Having a dedicated area/object/character as a connection to the real world justifies their use for delivering "real world" notifications in a non-intrusive way. Analogies can be drawn to Marti et al.'s work [16] on the use of animatronics for mobile phone notifications.

Avoiding Jump Scares. Visual and audio notifications that are spawned too close to the body may lead to undesired jump scares. Person talking (PAG) notifications caused many jump scares because of their sudden appearance close to the user and their life size. We suggest the spawning distance from the body of any notification must be customizable. Notifications closer to the body are also often perceived as more urgent, demanding of immediate attention. Undesirable experiences could also be reduced by timing the delivery of notifications to avoid interrupting any critical tasks.

Using Familiar Metaphors. To aid in the easy comprehension of VR notifications, abstraction based on real-world metaphors should be utilized when deemed fitting. Based on related work, abstraction is a popular choice in the design of ambient displays [2, 24] as well. In our studies, footstep sounds and the alternating left-right vibrations, as metaphors of people moving around, were well received by all. Similarly, all participants appreciated the similarity of the phone call notification to the interface of their phones.

Switching Context and Providing Details On-Demand. Notifications should be interactable and easily dismissible. Participants were not satisfied with simply receiving a notification. For example, they wanted to interact with the pop-up (POP) notifications to respond. Several also tried interacting with their virtual watches using the right controllers. Users should be able to switch context and easily make handling notifications their primary task in VR. Similar to other information displays, developers should also support the functionality of providing details on-demand. For example, revealing more details when users approach a notification, similar to semantic zoom.

8 Conclusion

We explored interruptibility and the design of notifications in virtual reality. We consulted VR experts to generate several design ideas for notifications in common interruption scenarios. We prototyped 5 notifications and conducted an empirical study across 6 output modalities in VR. The results show interesting effects of modality and notification type on reaction time, urgency, and understandability. Finally, combining our empirical results and participant feedback, we generated recommendations for VR notification designers and developers.

9 Acknowledgements

Special thanks to Dr. Simon Perrault, Meghna Natraj, Pratik Shah, Karan P. Singh, and Dr. Rafael Ballagas for their valuable feedback. We also thank our anonymous reviewers for helping to make this work better, as well as everyone who participated in our online survey, co-design exercise, and empirical studies for making this work possible.

References

[1] M. Adcock and S. Barrass. Cultivating Design Patterns for Auditory Displays. In *Proceedings of ICAD '04*, 2004.

[2] M. G. Ames and A. K. Dey. Description of Design Dimensions and Evaluation for Ambient Displays. Computer Science Division, University of California, 2002.

[3] A. Asif and S. Boll. Where to Turn My Car?: Comparison of a Tactile Display and a Conventional Car Navigation System Under High Load Condition. In *Proceedings of AutomotiveUI '10*, pp. 64–71, 2010. doi: 10.1145/1969773.1969786

[4] M. Azmandian, M. Hancock, H. Benko, E. Ofek, and A. D. Wilson. Haptic Retargeting: Dynamic Repurposing of Passive Haptics for Enhanced Virtual Reality Experiences. In *Proceedings of CHI '16*, pp. 1968–1979. ACM, 2016. doi: 10.1145/2858036.2858226

[5] P. Budhiraja, R. Sodhi, B. Jones, K. Karsch, B. Bailey, and D. Forsyth. Where's My Drink? Enabling Peripheral Real World Interactions While Using HMDs. *arXiv*, Feb. 2015. arXiv: 1502.04744.

[6] S. Chagué and C. Charbonnier. Real Virtuality: A Multi-user Immersive Platform Connecting Real and Virtual Worlds. In *Proceedings of VRIC '16*, pp. 4:1–4:3. ACM, 2016. doi: 10.1145/2927929.2927945

[7] M. Cidota, S. Lukosch, D. Datcu, and H. Lukosch. Workspace Awareness in Collaborative AR Using HMDs: A User Study Comparing Audio and Visual Notifications. In *Proceedings of AH '16*, pp. 3:1–3:8. ACM, 2016. doi: 10.1145/2875194.2875204

[8] E. Costanza, S. A. Inverso, E. Pavlov, R. Allen, and P. Maes. Eye-q: Eyeglass Peripheral Display for Subtle Intimate Notifications. In *Proceedings of MobileHCI '06*, pp. 211–218. ACM, 2006. doi: 10.1145/1152215.1152261

[9] K. Crommentuijn and F. Winberg. Designing Auditory Displays to Facilitate Object Localization in Virtual Haptic 3d Environments. In *Proceedings of Assets '06*, pp. 255–256. ACM, 2006. doi: 10.1145/1168987.1169042

[10] T. Grah, F. Epp, M. Wuchse, A. Meschtscherjakov, F. Gabler, A. Steinmetz, and M. Tscheligi. Dorsal Haptic Display: A Shape-changing Car Seat for Sensory Augmentation of Rear Obstacles. In *Proceedings of AutomotiveUI '15*, pp. 305–312. ACM, 2015. doi: 10.1145/2799250.2799281

[11] S. Hudson, J. Fogarty, C. Atkeson, D. Avrahami, J. Forlizzi, S. Kiesler, J. Lee, and J. Yang. Predicting Human Interruptibility with Sensors: A Wizard of Oz Feasibility Study. In *Proceedings of CHI '03*, pp. 257–264. ACM, New York, NY, USA, 2003. doi: 10.1145/642611.642657

[12] J. Häkkilä and S. Ronkainen. Dynamic Auditory Cues for Event Importance Level. In *Proceedings of ICAD '03*, pp. 233–237, 2003.

[13] N. Kern and B. Schiele. Context-aware Notification for Wearable Computing. In *Proceedings of ISWC '03*, pp. 223–230. IEEE, 2003.

[14] J. D. Lee, J. D. Hoffman, and E. Hayes. Collision Warning Design to Mitigate Driver Distraction. In *Proceedings of CHI '04*, pp. 65–72. ACM, 2004. doi: 10.1145/985692.985701

[15] A. Lucero and A. Vetek. NotifEye: Using Interactive Glasses to Deal with Notifications While Walking in Public. In *Proceedings of ACE '14*, pp. 17:1–17:10. ACM, 2014. doi: 10.1145/2663806.2663824

[16] S. Marti and C. Schmandt. Physical Embodiments for Mobile Communication Agents. In *Proceedings of UIST '05*, pp. 231–240. ACM, 2005. doi: 10.1145/1095034.1095073

[17] T. Matthews. Designing and Evaluating Glanceable Peripheral Displays. In *Proceedings of DIS '06*, pp. 343–345. ACM, 2006. doi: 10.1145/1142405.1142457

[18] D. S. McCrickard, C. M. Chewar, J. P. Somervell, and A. Ndiwalana. A Model for Notification Systems Evaluation—Assessing User Goals for Multitasking Activity. *ACM Trans. Comput.-Hum. Interact.*, 10(4):312–338, Dec. 2003. doi: 10.1145/966930.966933

[19] D. C. McFarlane and K. A. Latorella. The Scope and Importance of Human Interruption in Human-computer Interaction Design. *Hum.-Comput. Interact.*, 17(1):1–61, Mar. 2002. doi: 10.1207/S15327051HCI1701_1

[20] P. Milgram and F. Kishino. A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information and Systems*, E77-D(12):1321–1329, 1994.

[21] M. J. Muller. PICTIVE—an Exploration in Participatory Design. In *Proceedings of CHI '91*, pp. 225–231. ACM, 1991. doi: 10.1145/108844.108896

[22] K. Mäkelä, J. Hakulinen, and M. Turunen. The Use of Walking Sounds in Supporting Awareness. *Proceedings of ICAD '03*, 2003.

[23] T. Nukarinen, R. Raisamo, A. Farooq, G. Evreinov, and V. Surakka. Effects of Directional Haptic and Non-speech Audio Cues in a Cognitively Demanding Navigation Task. In *Proceedings of NordiCHI '14*, pp. 61–64. ACM, 2014. doi: 10.1145/2639189.2639231

[24] Z. Pousman and J. Stasko. A Taxonomy of Ambient Information Systems: Four Patterns of Design. In *Proceedings of AVI '06*, pp. 67–74. ACM, 2006. doi: 10.1145/1133265.1133277

[25] S. Razzaque. *Redirected Walking*. PhD Thesis, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, 2005.

[26] T. Roumen, S. T. Perrault, and S. Zhao. NotiRing: A Comparative Study of Notification Channels for Wearable Interactive Rings. In *Proceedings of CHI '15*, pp. 2497–2500. ACM, 2015. doi: 10.1145/2702123.2702350

[27] C. Schatzschneider, G. Bruder, and F. Steinicke. Who turned the clock? Effects of Manipulated Zeitgebers, Cognitive Load and Immersion on Time Estimation. *IEEE Transactions on Visualization and Computer Graphics*, 22(4):1387–1395, Apr. 2016. doi: 10.1109/TVCG.2016.2518137

[28] L. Shapira and D. Freedman. Reality Skins: Creating Immersive and Tactile Virtual Environments. In *Proceedings of ISMAR '16*, pp. 115–124. IEEE, Sept. 2016. doi: 10.1109/ISMAR.2016.23

[29] A. L. Simeone, E. Velloso, and H. Gellersen. Substitutional Reality: Using the Physical Environment to Design Virtual Reality Experiences. In *Proceedings of CHI '15*, pp. 3307–3316. ACM, 2015. doi: 10.1145/2702123.2702389

[30] S. Song, G. Noh, J. Yoo, I. Oakley, J. Cho, and A. Bianchi. Hot & Tight: Exploring Thermo and Squeeze Cues Recognition on Wrist Wearables. In *Proceedings of ISWC '15*, pp. 39–42. ACM, 2015. doi: 10.1145/2802083.2802092

[31] A. Stouffs. Interruptions As Multimodal Outputs: Which Are the Less Disruptive? In *Proceedings of ICMI '02*, pp. 479–482. IEEE Computer Society, 2002. doi: 10.1109/ICMI.2002.1167043

[32] E. A. Suma, Z. Lipps, S. Finkelstein, D. M. Krum, and M. Bolas. Impossible Spaces: Maximizing Natural Walking in Virtual Environments with Self-Overlapping Architecture. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):555–564, 2012. doi: 10.1109/TVCG.2012.47

[33] Q. Sun, L.-Y. Wei, and A. Kaufman. Mapping Virtual and Physical Reality. *ACM Trans. Graph.*, 35(4):64:1–64:12, 2016. doi: 10.1145/2897824.2925883

[34] P. Tarasewich, T. Bhimdi, and M. Dideles. Testing Visual Notification Cues on a Mobile Device. In *Proceedings of CHI EA '04*, pp. 1562–1562. ACM, New York, NY, USA, 2004. doi: 10.1145/985921.986138

[35] A. Telpaz, B. Rhindress, I. Zelman, and O. Tsimhoni. Haptic Seat for Automated Driving: Preparing the Driver to Take Control Effectively. In *Proceedings of AutomotiveUI '15*, pp. 23–30. ACM, 2015. doi: 10.1145/2799250.2799267

[36] M. Tomitsch, K. Kappel, A. Lehner, and T. Grechenig. Towards a Taxonomy for Ambient Information Systems. In *Proceedings of Pervasive '07*. Workshop: W9 - Ambient Information Systems, 2007.

[37] A. Voit, T. Machulla, D. Weber, V. Schwind, S. Schneegass, and N. Henze. Exploring Notifications in Smart Home Environments. In *Proceedings of MobileHCI '16 Adjunct*, pp. 942–947. ACM, 2016. doi: 10.1145/2957265.2962661

[38] A. Voit, D. Weber, and S. Schneegass. Towards Notifications in the Era of the Internet of Things. In *Proceedings of IoT '16*, pp. 173–174. ACM, 2016. doi: 10.1145/2991561.2998472

[39] D. Weber, S. Mayer, A. Voit, R. Ventura Fierro, and N. Henze. Design Guidelines for Notifications on Smart TVs. In *Proceedings of TVX '16*, pp. 13–24. ACM, 2016. doi: 10.1145/2932206.2932212

[40] B. G. Witmer, C. J. Jerome, and M. J. Singer. The Factor Structure of the Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 14(3):298–312, June 2005. doi: 10.1162/105474605323384654