Impact of Task on Attentional Tunneling in Handheld Augmented Reality

Conference Paper · May 2021

DOI: 10.1145/3411764.3445580

CITATIONS
20

READS
544

5 authors, including:

Brandon Syiem
The University of Sydney
11 PUBLICATIONS **56** CITATIONS
SEE PROFILE

Ryan Kelly
RMIT University
87 PUBLICATIONS **1,299** CITATIONS
SEE PROFILE

Jorge Goncalves
University of Melbourne
225 PUBLICATIONS **4,989** CITATIONS
SEE PROFILE

Eduardo Velloso
University of Melbourne
144 PUBLICATIONS **3,448** CITATIONS
SEE PROFILE

Impact of Task on Attentional Tunneling in Handheld Augmented Reality

Brandon Victor Syiem
Computing and Information Systems
University of Melbourne
Melbourne, VIC, Australia
bsyiem@student.unimelb.edu.au

Ryan M. Kelly
Computing and Information Systems
University of Melbourne
Melbourne, VIC, Australia
ryan.kelly@unimelb.edu.au

Jorge Goncalves
Computing and Information Systems
University of Melbourne
Melbourne, VIC, Australia
jorge.goncalves@unimelb.edu.au

Eduardo Velloso
Computing and Information Systems
University of Melbourne
Melbourne, VIC, Australia
eduardo.velloso@unimelb.edu.au

Tilman Dingler
Computing and Information Systems
University of Melbourne
Melbourne, VIC, Australia
tilman.dingler@unimelb.edu.au

ABSTRACT

Attentional tunneling describes a phenomenon in Augmented Reality (AR) where users excessively focus on virtual content while neglecting their physical surroundings. This leads to the concern that users could neglect hazardous situations when using AR applications. However, studies have often confounded the role of the virtual content with the role of the associated task in inducing attentional tunneling. In this paper, we disentangle the impact of the associated task and of the virtual content on the attentional tunneling effect by measuring reaction times to events in two user studies. We found that presenting virtual content did not significantly increase user reaction times to events, but adding a task to the content did. This work contributes towards our understanding of the attentional tunneling effect on handheld AR devices, and highlights the need to consider both task and context when evaluating AR application usage.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; **Mixed / augmented reality**; **Mobile devices**.

KEYWORDS

Attentional Tunneling, Augmented Reality, Mobile Devices

ACM Reference Format:
Brandon Victor Syiem, Ryan M. Kelly, Jorge Goncalves, Eduardo Velloso, and Tilman Dingler. 2021. Impact of Task on Attentional Tunneling in Handheld Augmented Reality. In *CHI '21: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, May 8 –13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3411764.3445580

1 INTRODUCTION

The use of mixed reality displays has become increasingly common in scenarios such as driving [4], aviation [32] and surgery [19, 24]. Although these displays can help users complete tasks more efficiently, studies have raised concerns about attentional issues that have potentially serious implications for user safety [17, 48, 51].

An attentional phenomenon that has been well-documented and associated with display designs is the *attentional tunneling effect*, which occurs when users focus excessively on virtual elements of a display while neglecting information presented in different channels, such as their physical surroundings [51]. Attentional tunneling leads to increased reaction times to events in information channels other than the channel that caused the effect [21, 29] or even missing these events altogether [51]. Previous work has proposed that the design of display technology is an important factor behind this effect, including head-up displays [22, 51], 3D perspective displays [21, 48] and also Augmented Reality (AR) displays [16–18]. As such, the conclusions of such studies often caution against the use of these display technologies in safety-critical situations [17].

The main finding of most attentional tunneling studies is that users performing a task while assisted by a display technology tend to miss events presented in channels outside of the display. To demonstrate this effect, studies have typically compared different types of display technologies—for example, head-up vs. head-down displays [22], 3D vs. 2D perspective displays [21], AR assisted vs. non-AR assisted displays [17]. They then report results on how one display type worsens the attentional tunneling effect over the other. However, a major limitation in studies that discuss attentional tunneling in relation to AR displays is that *they often confound the presence of the virtual content within the display with the additional task of scanning virtual content for information.* Taking this additional task into account is critical because it is well-established that users tend to miss events when engaged with a task, even though that event was plainly visible in their foveal view—an effect known as *inattentional blindness* [43, 46, 51].

Among the different categories of AR technologies, this paper focuses on applications presented on handheld Video See-Through (VST) AR devices, such as smartphones and tablets. Our decision to focus on handheld VST AR devices was based on the widespread adoption of handheld AR applications [2] and the potential issues of using AR applications while in motion [50, 55]. For example, the *Pokémon GO* AR game has had enormous success with the number

CHI '21, May 8 –13, 2021, Yokohama, Japan

of downloads reaching 1 billion [13], inspiring a whole genre of AR games [27]. However, the same game also brought AR under public scrutiny due to concerns related to user safety while walking, driving, and biking [9, 50]. These concerns stem from the observation that users pay little attention to their physical surroundings when engaged with the virtual elements of an AR application, a phenomenon reminiscent of the attentional tunneling effect.

AR content is registered with the physical world by definition [10]. Previous studies have reported conflicting results in terms of the relationship between AR and attentional tunneling. On the one hand, studies have suggested that imagery that conforms to the physical world tends to negate or mitigate the attentional tunneling effect [30, 31]. On the other, studies have put forward concerns related to an excessive focus on virtual content in AR applications [17, 18]. This suggests a critical need for teasing apart the potential causes of attentional tunneling in AR. Understanding the causes of attentional tunneling in AR, alongside situations in which it does and does not occur, can support designers in identifying situations in which AR can be most successfully applied. In this paper, we question whether the attentional tunneling effect is indeed caused by the virtual content itself or by the task the user is engaged with while using the AR application. We hypothesise that the cause of the attentional tunneling effect is the presence of the additional task of scanning virtual elements for information, as opposed to the virtual content itself as suggested by previous work.

We conducted two studies to understand whether virtual elements in handheld AR displays cause the attentional tunneling effect and how the presence of a task related to the virtual elements impacts the phenomenon. We operationalise attention in terms of user reaction times in detecting events [26] given that increased reaction times have been associated with the attentional tunneling effect [21, 51]. In both experiments, we measured reaction times to events seen through a display while manipulating the presence of virtual content and associated task. The first experiment was administered through an AR-enabled smartphone. The second experiment was a web-based simulation of the first study that aimed to control for the peripheral visibility of events outside the frame of the smartphone used in the first experiment. This setup mirrors real-world use cases of AR in which the user has no visibility outside of the display, such as in vehicle navigation [4] or remote manipulation of robots [40].

Overall, we found no significant increase in reaction times with the mere presence of virtual content in an AR application, but the effect was significant when users were asked to accomplish a task related to this content. This finding suggests that the 'attentional sink' into virtual content within an AR application is a result of users scanning virtual content for information to complete a task, rather than the presence of virtual content itself. In the first study, we also found that users were faster at reacting to stimuli originated from the physical world than to visually similar stimuli that were only rendered on the phone. We hypothesise that this difference is due to the peripheral visibility of such stimuli. We confirmed this hypothesis in the second study in which we controlled for the effect of peripheral visibility. This suggests that the smaller field of view of handheld AR-applications may ameliorate the attentional tunneling effect, provided that users have visibility of events happening outside the AR content.

Our results contribute towards a better understanding of the attentional tunneling effect in handheld AR applications by dispelling doubts expressed in previous work [12, 17, 18] that virtual content in AR applications causes the attentional tunneling effect. The results also highlight the need to evaluate AR applications for user risks related to attentional issues based on the task the application is used for, not just the virtual content itself.

2 RELATED WORK

Our coverage of the related work is divided into two subsections. We first describe the attentional tunneling effect and its reported causes. We then focus on attentional issues in AR displays.

2.1 Attentional Tunneling

The attentional tunneling effect was formally defined by Wickens et al. [51, p. 182] as:

> "The allocation of attention to a particular channel of information, diagnostic hypothesis, or task goal, for a duration that is longer than optimal, given the expected cost of neglecting events on other channels, failing to consider other hypotheses, or failing to perform other tasks."

The effect has been reported to have a number of different causes such as increased cognitive load [11, 38, 52], task priority and complexity [41], conversation [8], operator fatigue [33], the priming effect [29] and the novelty and infatuation with new technology [51]. It has also been reported to be caused by different display designs such as HUDs [22, 51], AR displays [16, 17] and 3D perpective displays [21, 48].

One of the first studies that reported the attentional tunneling effect was conducted by Fischer et al. [22]. The study compared pilot performance in an aircraft cockpit simulator when using a traditional head-down display (HDD) as compared to a head-up display (HUD) to present flight relevant information (altitude, flight path, etc). The HUD presented information superimposed on the pilot's view of the outside world, while the HDD presented information on a separate screen placed below the outside world scene, similar to a car dashboard. Though participants' performances were better using the HUD, the authors reported that pilots tended to miss more unexpected events on the view of the outside world when using the HUD as compared to the HDD. This finding was surprising because the HUD preserved the outside world in foveal view, while using the HDD required users to shift their gaze between the two contexts.

An explanation as to why this effect is prevalent on HUD displays is related to how users perceptually group information [31]. HUD elements are presented on a separate depth layer to the background content (outside world view) and hence behave differently from the outside world view—for example, HUD components are stationary even when the outside view changes with respect to user motion. This prompts users to group the HUD elements and outside view into separate planes of information, hindering their ability to divide their attention between the two planes [51]. A solution to this is to 'symbolically link' the HUD components to the view of the outside world, i.e., using virtual components in the HUD that are registered

with the outside world view and respond to changes in user positions accordingly [31]. This method has been shown to mitigate or completely eliminate the attentional tunneling effect [30, 31].

Attentional tunneling not only causes users to miss unexpected events, but it causes significant delays in reaction times to events not presented in the same channel. Fadden et al. demonstrated a significant 4-second delay to reaction times of pilots to unexpected events when using 3D content displayed on a HUD, exemplifying the attentional tunneling effect [21]. Kortschot et al. induced the attentional tunneling effect on participants to study if they would miss events on different channels [29]. They grouped participants into 'tunneled' and 'non-tunneled' groups whereby the 'tunneled' group were primed [25] to expect events in a specific location on a 2D screen. The final event was placed outside of this location and the researchers observed a significant increase of approximately 3 seconds in reaction time to the final event for the 'tunneled' group when compared to the 'non-tunneled' group.

Aside from the initial reports of HUDs causing the attentional tunneling effect [22], evidence also suggests that 3D content could worsen the effect. Fadden et al. compared reaction times of pilots when using 2D versus 3D content presented on HUDs and concluded that 3D perspective displays increased the attentional tunneling effect [21]. Studies unrelated to attentional tunneling have also reported increased reaction times when comparing 3D egocentric and 2D exocentric displays [48]. However, the observation that 3D content causes the effect has not been consistently observed in other studies. For example, Laswell et al. reported that pilots were able to detect a greater number of unexpected events in the outside world view when using a 3D conformal display as opposed to a conventional 2D non-conformal display [30]. Olmos et al. also reported that that reaction times to unexpected events were significantly higher in a 2D co-planar display when compared to a 3D exo-centric (or bird's eye perspective) display [34].

Attentional tunneling has also been associated with the type of information presented on the display. In their study in an aviation context, Wickens et al. explored two displays presenting different flight-relevant information in conventional head-down 2D monitors rather than HUDs [51]. They compared the attentional tunneling effect between a system that previewed outside hazards (Synthetic Vision System or SV [37]) to a 3D flight navigation system called Highway in the Sky (HITS) [3]. The effect was observed to be more prominent on the HITS display when compared to the SV display. The authors attributed this to the different design goals of each display: while the SV was designed to inform users of the state of the outside world, the HITS display was designed to present more involving information that implicitly or explicitly guided the user [51]. The authors also attributed the increased focus on the HITS display to its 3D presentation.

Dixon et al. explored the attentional tunneling effect in a controlled experiment related to AR-assisted endoscopic navigation [16, 17]. They reported that while performance was improved by the AR-assisted display when compared to a display with no AR content, the participants failed to detect foreign objects while performing the endoscopic test. The researchers concluded that AR displays did indeed cause attentional tunneling [16, 17] and proceeded to caution against the improper use of AR for medical applications.

2.2 Augmented Reality and User Attention

The primary characteristic of AR technology is its ability to present virtual content overlaid and registered to the physical world [35]. This enables users to move through and interact with their surrounding physical environment while also having access to additional virtual content. As such, the promise of AR technology is that it should not hinder the actions in which people normally engage in the physical world, including object manipulation and social interactions. This has led to a number of systems and studies dedicated to the use of AR in public settings [14, 39], collaborative tasks [12, 36], classroom learning environments [20] and games in public spaces [27], among others.

However, following the release of the *Pokémon GO* AR game [7], a number of studies have discussed the dangers of using AR applications while walking, driving, or cycling [9, 50]. Although *Pokémon GO* preserved the users' physical surroundings within view, users nevertheless paid little attention to their physical surroundings while playing the game. Chang et al. also recognised this concern in the use of AR in museum settings and described the limitation of mobile AR-guide systems as "the inability to balance a visitor's attention distribution between the virtual space and the physical scenes" [15, p. 186].

The lack of attention paid to physical elements in public AR applications can also hinder the user experience in other contexts [47]. Consider, for example, an AR educational application used to augment the learning environment with relevant virtual content. If students pay too much attention to these virtual augmentations and neglect crucial elements in their physical surroundings, such as the teacher or peers, the purpose of the application—i.e. to support the intended learning objectives—would be defeated.

Although studies have discussed the relationship between AR and the attentional tunneling effect, they have done so mostly in passing and without a formal study exploring the attentional tunneling phenomenon in AR applications [12, 15, 18]. The few studies that did explore the attentional tunneling effect in AR have done so without controlling for the effects of the additional task of scanning for relevant information imposed by the virtual elements [16, 17]. This scanning task may act as a potential confound, given that engagement with a task is known to cause attentional issues [41, 46]. The addition of a secondary task has also been known to reduce performance in visual stimulus response task in AR applications [28].

3 STUDY 1

3.1 Method

In this study, we aimed to disentangle the impact of the virtual content and the task that relates to it on the attentional tunneling effect in the context of handheld VST AR applications. As such, we conducted a lab study using a handheld AR device in which we measured reaction times while manipulating the content and tasks in each condition. All of our materials and procedures received approval from our institution's ethical review panel.

3.1.1 Experiment Design. We investigated the causes of attentional tunneling by measuring the effects of content and task on the reaction time in response to events that happened on a handheld AR display and in the physical world (as seen through the display). We operationalised the attention tunneling effect by measuring the time

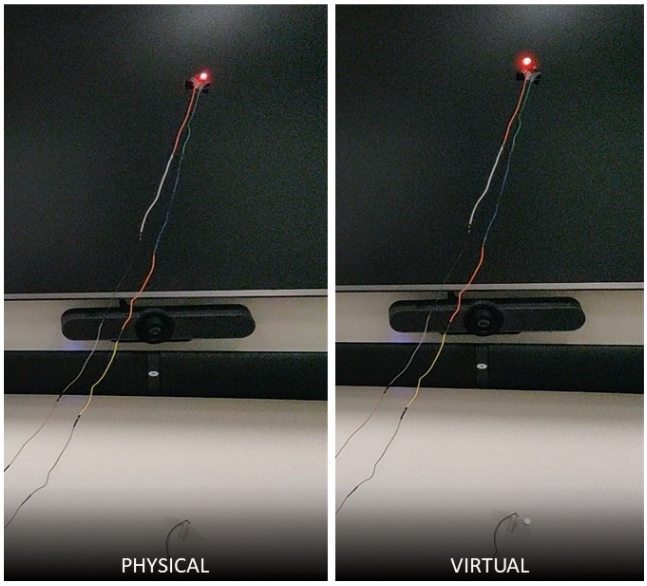

Figure 1: Screenshots of the camera feed observing the physical LED and of the virtual LED rendered over the camera feed. The virtual LEDs were designed to be as similar as possible to the physical ones.

to react to an event [26] as previous works have demonstrated that this effect leads to increased reaction times [21, 29].

We also recorded the number of missed events for each condition given that prior work has reported that the attentional tunneling effect causes users to miss events [30, 34, 51]. However, past works have also shown that the attentional tunneling effect does not necessarily cause users to miss events [29, 53] and as such, we primarily focus our discussion on the reaction time results.

In our experimental task, participants were asked to press a button whenever they detected the onset of the event of an LED turning on, as seen through the AR display. Only one LED would be active at any given instance. The LEDs turned on at random intervals between 4 and 10 seconds since the last LED turned on so that participants could not anticipate the event. LEDs would remain on until the button was pressed or the next LED turned on after 4-10 seconds. The 4–10 seconds interval was based on pilot testing, which indicated that the interval gave enough time for the user to react to events without allowing the user's attention to drift. We measured reaction time as the difference in milliseconds between the LED being lit up and the user pressing the button.

We employed a 2x3 repeated measures design with Event Type (2 levels) and Content (3 levels) as independent variables. The first independent variable (IV), Event Type, corresponded to whether the LED was virtually rendered on the screen or physically lit by an Arduino. We controlled for the appearance and position of the LEDs by designing the stimuli to be near-identical in terms of their size, brightness, colour, and position on the screen (see Figure 1). This IV allowed us to compare the reaction times for stimuli that happen in the *same channel* as the cause of the effect (Event Type:Virtual) to stimuli that happen in a *separate channel* as the cause of the effect (Event Type:Physical). According to related work (e.g. [30, 31]), the attentional tunneling phenomenon only affects reaction times to events presented in a separate channel to that which caused the attentional tunneling phenomenon. Hence, if virtual content in an AR display causes the attentional tunneling effect, we would expect to see larger reaction times (i.e., slower reaction) for Event Type:Physical as compared to Event Type:Virtual, but if the effect is caused by the task, then there should not be any difference in these reaction times.

The second independent variable was the Content shown on the AR display. As a baseline (Content:None), no additional content other than the feed of the camera and the stimuli (i.e. LEDs) were shown on the screen. In Content:Blocks, 5 blue blocks rendered in AR moved around the registered space while passing a red sphere between them. Content:Blocks + Task was used to introduce a task-based condition and was visually the same as Content:Blocks, but participants were asked to count the number of times the blocks passed the sphere to one another. This task was inspired by Simons and Chabris' study on inattentional blindness [46], in which participants counted the number of times a basketball was passed amongst a team of players. This allowed us to compare conditions with a task (Content:Blocks + Task) and without a task (Content:None and Content:Blocks) to determine how the task associated to the virtual content affected the attentional tunneling effect. It also enabled us to compare conditions without tasks but differing in the presence of virtual content (Content:None and Content:Blocks) to test whether merely presenting virtual content on a AR display causes the attentional tunneling effect, as suggested in previous literature [16–18].

The independent variables we consider and their levels are summarised as follows:

(1) Event Type:
Virtual - users were asked to press a button when a virtual LED was lit on the screen;
Physical - users were asked to press a button when a physical LED was lit in the physical world.
(2) Content:
None - only the feed from the camera was shown on the mobile device;
Blocks - virtual moving blocks, passing a red ball between them, were overlaid onto and registered to the feed from the camera;
Blocks + Task - the same as in *Blocks*, but with users asked to count the number of ball passes between the blocks.

Following our ANOVA, we test for equivalence between conditions where we expect no significant effects using the Two One-Sided T-test (TOST) procedure [44, 45]. This approach enables us to test if an effect is statistically equivalent to zero based on an upper and lower bound. We chose the lower bound of effect size to be equal to -0.2 and the upper bound to be +0.2 to correspond to the Cohen's D 'negligible' effect size. The test reverses the roles of the null and alternative hypothesis i.e., the null hypothesis states that the effect lies outside the range of the lower or upper bound, while the alternative hypothesis indicates that the effect lies within the defined range.

Based on the evidence presented in past literature which indicates that engagement with a task causes inattentional blindness [46] and attentional tunneling [41], we put forward the following hypothesis:

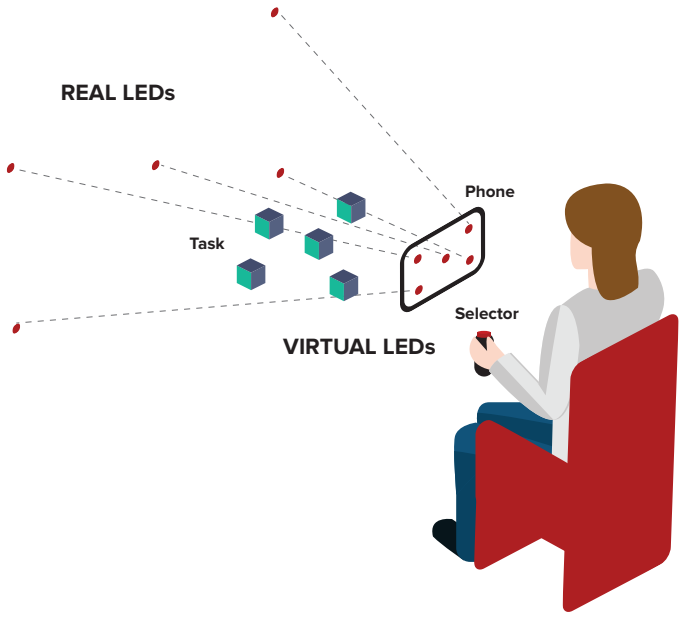

Figure 2: An illustration of our study setup depicting a user holding a button (Selector) to react to LED events viewed through a handheld AR device. LEDs were either real or virtual. Virtual content in the form of floating and moving blocks was rendered on the device. The user was asked to perform a task related to the blocks depending on the experiment condition.

Hypothesis 1. *Reaction times to events (within the same event type) in a handheld AR application will be significantly longer when virtual AR content is accompanied by a task.*

We expected significant differences in reaction times between conditions with (Content:Blocks + Task) and without (Content:None and Content:Blocks) a task, irrespective of Event Type.

In contrast, we did not expect significant differences in reaction times within event types between the conditions that present no virtual content (Content: None), to conditions with virtual content and no associated task (Content: Blocks). While past literature has demonstrated that AR displays can cause the attentional tunneling effect [16, 17], they have done so in the presence of a potential confound, i.e., the secondary task of scanning virtual elements for information. There is also evidence to suggest that conformal or registered content in a digital display, as with the case of AR displays [10], can mitigate or even negate the attentional tunneling effect [30, 31]. Therefore, our second hypothesis was:

Hypothesis 2. *There will be no significant difference in reaction times to events (within the same event type) between conditions that present no virtual content and conditions that do present virtual content with no associated task.*

Past literature has also reported that the attentional tunneling effect does not impact the detection of events presented in the same channel which causes the effect [30, 31]. This means that we should see a significant increase in reaction times to physical events as opposed to virtual events. This is because prior work has expressed how virtual content within an AR environment draws excessive user attention [17, 18]. However, because we hypothesise that the presence of the secondary task is causing the attentional tunneling effect, we do not expect to see faster reaction to virtual events within the same Content level.

3.1.2 Experiment Setup. Figure 2 illustrates our experimental setup. The system consisted of a OnePlus 6 smartphone with an Android 8.1 operating system and two Arduino Mega2560 microcontrollers [1]. The first Arduino was connected to five *physical* LEDs and a single button that was used by participants to react to an LED turning on. The second Arduino was connected to a single button that was used by participants to react to *virtually rendered* LEDs presented on the smartphone. The LEDs (both physical and virtual) were placed in the top-left, top-right, bottom-left, bottom-right and center of the smartphone's screen view. The smartphone was connected to a single Arduino at a time via a USB type A/B cable for serial communication. A Google ARCore-enabled [6] Unity application ran on the smartphone to display virtual objects rendered onto the smartphone's back camera image feed.

Application: The Unity application controlled the Content presented on the smartphone while also managing the communication with the Arduino. Depending on the level of Event Type and Content, the application presented different scenes: the camera feed with no additional content, the camera feed with virtual LEDs, the camera feed with blocks content, or the camera feed with the blocks content and virtual LEDs. Figure 3 shows the participant's view of the smartphone with all virtual content present—blocks, ball, and virtual LEDs.

In all conditions, the application waited for a random amount of time (between 4 to 10 seconds) and sent out a message to the Arduino

Figure 3: View of smartphone as seen by participants which shows the setup with blocks passing a red sphere and virtual LEDs. We arranged the furniture of the space to resemble a living room to increase the external validity of the task.

to start the timer and turn on an LED. The Arduino then took over the measurement and communication of the reaction times based on button press events. Pressing the button when an LED is in the 'on' state records the time elapsed between the onset of the event and the button press. Pressing the button when no LED is on (i.e., a false positive) was also recorded. Not reacting to an LED event before another LED event triggered was considered as a 'missed event' and was also recorded. In conditions where the blocks were rendered, the application also recorded the number of times the sphere was passed between the blocks.

3.1.3 Participants. Twenty participants (men = 14, women = 6) aged between the ages of 18 and 32 ($M = 25.2, SD = 4.2$) took part in the experiment. Participants were given a $10 gift voucher for participation.

Each participant performed the reaction test in all 6 different conditions (within-subject design) and we collected 25 reaction time data points for each condition.

3.1.4 Procedure. The experiment took place in a controlled laboratory environment, styled as a living room to simulate a natural setting where users are likely to use an AR application. We ensured that some of the furniture was in view as visual clutter for the background to increase the task difficulty and ecological validity. The furniture arrangement was identical for all participants. We asked participants to sit on a couch, in front of a stand with our smartphone device and wear a head-mounted eye tracker (Tobii Pro Glasses 2) to record eye movements. The eye tracker allowed us to verify whether participants were paying attention to the task during the experiment. In particular, we were interested in verifying whether they looked at the physical LEDs directly or through the phone.

All participants read a plain language statement and provided written informed consent. After calibrating the eye tracker, we explained the task, informing participants that they should press the button whenever any LED lit up on the screen (either virtual or physical). In conditions involving the additional task (Content: Blocks + Task), we informed participants that in addition to reacting to the LEDs, they should count the number of times the sphere was passed between the virtual blocks. To motivate participants to count accurately, we told them that there would be a surprise reward (which participants later found out to be chocolate) if they recorded the number of passes correctly. After the experiment, we conducted a short interview in which we asked participants whether they had any difficulties using the button to react to the events. We also asked participants which task they found the most difficult for each event; which event type was easier to detect; and why was it easier to detect.

3.1.5 Pre-processing. Before testing our hypotheses, we analysed the data to identify any instances of lack of adherence to the experimental protocol. No participants were excluded from this experiment. The experiment comprised of a total of 120 samples (6 conditions x 20 participants).

To ensure that participants adhered to the instructions of the study, both in terms of looking at the device (i.e. not outside the device to detect physical events) and also in terms of engaging with the counting task in the task based conditions, we analysed the error rate for the counting task and manually examined the eye tracker recordings. The eye tracking data indicated that all participants maintained their gaze on the smartphone screen throughout the experiment. The counting task was also successful in engaging participants, with an average error rate of 0.9 percent.

We also limit our discussion to reaction times as the number of recorded missed events in the study were insufficient for a statistical test with a total of 44 recorded missed events (0.001%) out of 3000 events.

3.2 Results

Our main finding was that participants exhibited significantly larger reaction times in conditions with a task (Content: Blocks + Task)

compared to conditions without tasks (Content: Blocks and Content: None). This effect held for both virtual and physical events. However, we found significantly larger reaction times to Event Type: Virtual when compared to Event Type: Physical for all levels of Content. This supports our hypothesis that task associated with virtual content within an AR application causes the attentional tunneling effect.

Additionally, we found that the observed effect between conditions with virtual content and no task (Content: Blocks) and conditions without virtual content Content: None were statistically equivalent to zero for both Event Type: Physical and Event Type: Virtual. This suggests that solely presenting virtual content without an associated task in a VST AR application does not induce attentional tunneling.

3.2.1 Quantitative Results. This section details the results of our analysis on participants' reaction times to events based on Event Type and Content.

Table 1: Summary Statistics for Sample Reaction Times grouped by Event Type and Content in Study 1.

Event Type	Content	N	Mean(ms)	SD(ms)
Physical	None	20	580	131
Physical	Blocks	20	585	144
Physical	Blocks + Task	20	691	199
Virtual	None	20	728	135
Virtual	Blocks	20	781	149
Virtual	Blocks + Task	20	945	200

Reaction Times. The summary statistics of reaction times grouped by Content and Event Type are detailed in Table 1. A Shapiro-Wilk test of the data grouped by Event Type and Content indicated a deviation from a normal distribution. We log-transformed the data and ran the Shapiro-Wilk test again, which no longer indicated a substantial deviation from a normal distribution. Levene's test did not show any violation of homogeneity of variances ($F_{5,114} = 0.33$, $p = 0.89$).

We therefore used a two-way repeated measures ANOVA to determine if there was any significant effect of Content, Event Type and their interaction on participants' reaction times. We found significant main effects of Event Type ($F_{1,19} = 79.94$, $p = 3.1E-8$, $partial\ \eta^2 = 0.81$) and Content ($F_{2,38} = 21.0$, $p = 7.1E-7$, $partial\ \eta^2 = 0.53$) on the reaction time. We found no significant interaction between the effects of Event and Content on participant reaction times.

We ran further analyses to determine the effect of Content on each level of Event Type using Bonferroni-corrected pairwise t-tests. The analysis revealed significantly longer reaction times for conditions that involved an additional task (Content: Blocks + Task) when compared to conditions without it (Content: None and Content: Blocks) for both virtual and physical events. This suggests that the presence of a task does induce the attentional tunneling effect in handheld AR applications. The significant results of the pairwise t-tests are detailed in Table 2. Figure 4 shows a boxplot indicating the significant differences of Content at each level of Event Type.

The pairwise t-tests did not reveal any significant difference in reaction times to both physical and virtual events between conditions

Table 2: Results of the pairwise t-test for the effect of Content on each level of Event Type for Study 1.

Event Type	Content 1	Content 2	t(19)	p adjusted	Cohen's d
Physical	None	Blocks + Task	-3.55	$6.0E-3$	0.79
Physical	Blocks	Blocks + Task	-3.17	$1.5E-2$	0.71
Virtual	None	Blocks + Task	-6.04	$2.5E-5$	1.35
Virtual	Blocks	Blocks + Task	-4.69	$4.8E-4$	1.1

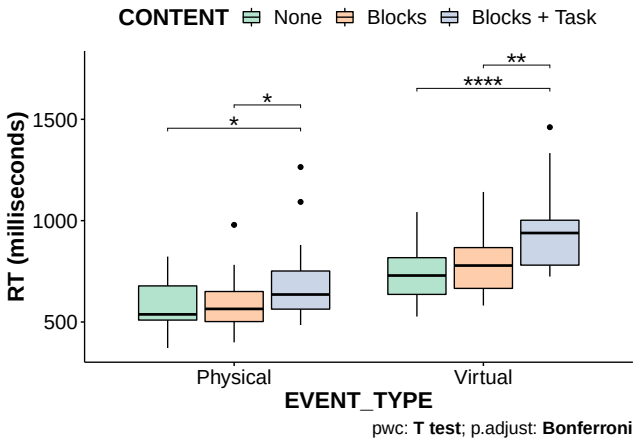

Figure 4: Effect of Content type on each Event Type level in study 1. The boxplot indicates that there was a significant increase in reaction time to both virtual and physical events when comparing conditions without (Content:None and Content:Blocks) and with (Content:Blocks + Task) a task.

with and without virtual content in the absence of the associated task. We therefore ran equivalence tests using the TOST procedure on conditions where no significant difference was detected in the pairwise t-test, i.e., between conditions with Content: None and Content: Blocks for both levels of Event Type. The analysis revealed that the observed effect between the conditions Content: None and Content: Blocks for Event Type: Physical was statistically equivalent to zero ($t_{38} = 2.66, p = 5.7E-3$). The analysis also revealed that the observed effect between the conditions Content: None and Content: Blocks for Event Type: Virtual was statistically equivalent to zero ($t_{38} = 2.20, p = 1.6E-2$).

Next, we analysed the effect of Event Type on participant reaction time for each level of Content using pairwise t-tests with Bonferroni corrections. The results of the tests are detailed in Table 3. The tests revealed that reaction times were larger for virtual events over physical events for every level of Content. This indicates that participants were significantly faster in reacting to the onset of the LEDs when they were presented in physical form, rather than as virtual stimuli on the smartphone. However, based on the definition of attentional tunneling, if attentional tunneling was caused by virtual content, participants should have reacted faster to events in the same channel [30, 51] i.e., virtual LEDs. This result therefore questions whether virtual content in AR applications can solely cause the attentional tunneling effect.

Table 3: Results of the pairwise t-test for the effect of EVENT on each level of CONTENT for study 1.

CONTENT	EVENT 1	EVENT 2	t(19)	p adjusted	Cohen's d
None	Physical	Virtual	-5.78	$1.5E{-}5$	1.29
Blocks	Physical	Virtual	-6.67	$2.2E{-}6$	1.49
Blocks + Task	Physical	Virtual	-7.61	$3.5E{-}7$	1.70

3.2.2 Qualitative Results. We used the interview data to explore factors that may have affected participant reaction times. To this end, we asked all participants if they had any difficulties pressing the buttons to react to the events. All participants confirmed that they had no difficulties in pressing the button to react to both the physical and virtual LED event.

We asked participants if they had prior experience with AR technology to consider any effects of familiarity with the technology on their reaction times. Nine participants mentioned that they had used AR applications before, mostly quoting *Pokémon GO*. However, all participants confirmed that they had no extensive experience with AR technology.

To explore factors that may have influenced reaction times and caused attentional tunneling, we asked participants to describe which conditions made it most difficult for them to react to the LED events. Participants unanimously agreed that CONTENT: BLOCKS + TASK was the most difficult and no participant mentioned noticing any difference in difficulty between CONTENT: BLOCKS and CONTENT: NONE. With respect to EVENT TYPE, the majority of participants (14/20) mentioned that the physical LEDs were easier to react to, disagreeing with past literature [31] that virtual events were easier to react to as participants would be 'tunneled' into the virtual content. One participant mentioned that there was no difference in difficulty when reacting to physical or virtual events while the remaining five participants mentioned that the virtual LEDs were easier to detect than the physical LEDs.

Thirteen out of the 14 participants who mentioned it was easier to detect the physical LEDs explained that it was easier because they could see them "outside of the device" or in their peripheral vision. One participant also added that they could see the reflection of light projected by the physical LED on nearby surfaces. This suggest that the smaller field of view of the smartphone, which accommodates for a larger peripheral view of the physical surroundings, may help in the detection of physical events, thereby providing a countermeasure for the attentional tunneling effect in handheld AR applications.

The five participants who mentioned that the virtual LEDs were easier to detect gave examples only related to the condition with CONTENT: Blocks + Task, claiming that it was easier to switch between looking at the sphere and scanning the virtual LEDs.

4 STUDY 2

4.1 Method

The first study was administered through a handheld AR device to preserve ecological validity of the results. However, the relatively small field of view of the device enabled participants to view physical events 'outside' the frame of the device, which resulted in faster reaction times to physical events (Section 3.2). To control for the peripheral visibility of physical events, we conducted a second study via a web-application that simulated the exact scene as seen through the handheld AR device but without any 'outside' view of the physical events.

The web-application setup also enabled us to explore how task impacts the attentional tunneling effect in desktop-based AR applications. While not as commonplace as handheld AR applications, desktop-based AR is used in applications such as navigation [4] and remote robot manipulation [40] where attentional issues may carry risks towards the user or the equipment. For example, Yeh and Wickens [54] demonstrated how the addition of virtual cues on a 2D display caused the attentional tunneling effect during target detection for unmanned aerial vehicles.

4.1.1 Experiment Design. The design of this experiment was aimed at understanding the effect of peripheral visibility of physical events on participant reaction times. As such, the design mimics the conditions and scene of the AR-based setup but without peripheral visibility of physical events outside the view of the display. We hypothesised that by removing this peripheral visibility, the effect of event type would disappear because users can no longer see the onset of LEDs outside the display or in their peripheral vision. We employed the same set of IVs (EVENT TYPE and CONTENT) and statistical procedures (2x3 ANOVA and TOST equivalence procedure) used in the AR-based setup (Section 3.1.1). This enabled us to determine if there is any significant difference in reaction time based on EVENT TYPE without the influence of peripheral visibility of physical events.

The web-based simulation differed from the AR-based setup (Study 1) in its presentation of the scene and events. While the AR-based setup used a handheld AR device and presented the camera feed with (or without) registered virtual content on the display, the web-based simulation displays an image of the exact scene as seen through the smartphone in the AR-based setup as the background image of the web-application. Displaying the scene as a background image ensured that the LEDs could only been seen on the display. The events in this case simulate an LED turning on by swapping the background image to one where an LED is switched on. Participants can react to an event by left clicking the mouse. A background image with a random LED in the on state would be displayed every 4-10 seconds and remain active until the user clicks the mouse or until another background image with a switched on LED is selected to be displayed. The IV, EVENT TYPE, corresponded to whether the background images were screenshots with the physical LEDs or the virtual LEDs as seen through the AR device in the AR-based setup. The second IV determined the CONTENT overlaying the background image in the web-application. CONTENT had the same three levels (None, Blocks and Blocks + Task) and functions as in the AR-based setup (described in section 3.1.1).

Since both EVENT TYPES (Physical and Virtual) cannot be seen outside the frame of the display on a users' peripheral view, we expect to see no significant difference in reaction times between EVENT TYPE: Physical and EVENT TYPE: Virtual within the same CONTENT level.

4.1.2 Experiment Setup. The web-based simulation was designed as a Unity WebGL application and mirrored the scene and functions of the AR-based applications (described in Section 3.1.2). The WebGL application presented the participants with the camera feed of the AR device used in the AR-based setup as a background image

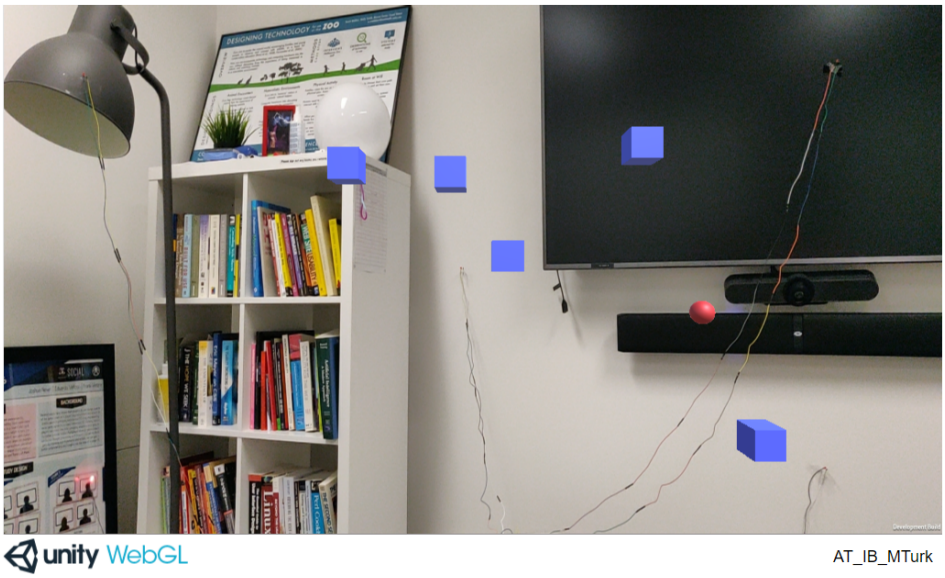

Figure 5: Image of the WebGL application running on a web browser window used in study 2. The application is presenting a background image with physical LEDs and is overlaid with virtual blocks.

(living room setup as seen in Figure 3). The application window was set to 27.1cm X 15.24cm (to maintain the aspect ratio of 16:9 of the background images). To trigger an event (switching on an LED), the application swapped the background image to one in which the required LED was in the 'on' state. Events were triggered in random intervals of 4 to 10 seconds, identical to the setup in study 1.

The Event Type variable determined which background images were used for a specific condition, i.e., background images capturing physical LEDs or background images with virtual LEDs. Each Event Type (Physical and Virtual) consisted of 6 different background images for a total of 12 images. For each Event Type, one image showed the background with all the LEDs turned off (base) and the other 5 images showed each of one LED in the on state.

The Content variable determined whether to render additional virtual elements over the background. In conditions with Content: None, no additional virtual elements were rendered, while conditions with Content: Blocks and Content: Block + Task, virtual blocks passing a virtual sphere were rendered over the background image.

Participants reacted to events by clicking with the left mouse button. The time taken between an event triggering (background LED turned on) and the time it took the participant to left-click the mouse was calculated and a WebRequest was made to our server to store the reaction time data to a persistent file. Clicking the mouse button when no LED was on was also recorded. Not reacting to an event before another event triggers was considered as a 'missed event', which was also recorded.

Figure 5 shows a screenshot of the WebGL application. The application used screenshot images depicting the camera feed from Study 1 and displayed these as background images. Each Event Type (Physical and Virtual) consisted of 6 different background images for a total of 12 images. For each Event Type, one image showed the background with all the LEDs turned off (base) and the other 5 images with each of one LED in the on state. To trigger an event (i.e., turn on an LED), the application would swap out the current background image for another image with the required LED in the on state.

4.1.3 Participants. Forty-eight participants (men = 24, women = 24) aged between the ages of 21 and 57 ($M = 33.3, SD = 8.9$) took part in the experiment. Participants were recruited via Amazon Mechanical Turk and were compensated with $20 for participating.

Each participant performed the reaction test in all 6 different conditions (within-subject design) and we collected 25 reaction time data points for each condition.

4.1.4 Procedure. All participants were directed (via hyperlink) to a Plain Language Statement and a Consent Form. Participants consented by selecting a checkbox on the Amazon MTurk survey form. Participants were then directed to our web-application via another hyperlink and could complete the experimental tasks.

After completing the tasks, participants filled out a short questionnaire hosted on Amazon MTurk, to gauge their experience with AR games and to investigate their opinions about each condition.

4.1.5 Pre-processing. Before testing our hypotheses, we analysed the data to identify any instances of lack of adherence to the experimental protocol. Three participants (men = 1, women = 2) were excluded from this experiment due to excessive mouse clicks (i.e., false positives), which resulted in unnaturally fast reaction times to events being recorded. The experiment comprised of a total of 270 samples (6 conditions x 45 participants).

To ensure that participants were engaged with the study, instructions presented on the web-based application clearly informed participants that reacting to the events and correctly performing the counting task (in the task based condition) was crucial for their attempt to be registered. Despite this, we observed an average error

rate in the counting task of approximately 16.7%, much larger than the 0.9% average counting error rate in the first experiment.

Similar to Study 1, we limit our discussion to reaction times as the number of recorded missed events in the study was insufficient for a statistical test with a total of 165 recorded missed events (0.02%) out of 6750 total events.

4.2 Results

Our main finding was that the observed effect between Event Type: Physical and Event Type: Virtual for conditions Content: Blocks and Content: Blocks + Task) were statistically equivalent to zero. This indicates that peripheral visibility of physical LEDs in Study 1 reduced the reaction times of participants to physical events.

We also found a significant difference in reaction times between conditions with tasks (Content: Blocks + Task) and without tasks (Content: Blocks and Content: None) for Event Type: Physical. This suggests that task associated with virtual content within an desktop-based AR application causes the attentional tunneling effect.

4.2.1 Quantitative Results. This section details the results of our analysis on user reaction times to events based on Event Type and Content.

Table 4: Summary Statistics for Sample Reaction Times grouped by event and content in Study 2.

Event Type	Content	N	Mean(ms)	SD(ms)
Physical	None	45	1116	613
Physical	Blocks	45	1191	628
Physical	Blocks + Task	45	1276	554
Virtual	None	45	1181	597
Virtual	Blocks	45	1145	522
Virtual	Blocks + Task	45	1221	545

Reaction Times. The summary statistics of reaction times grouped by Content and Event Type are detailed in Table 4. A Shapiro-Wilk test of the data grouped by Event Type and Content indicated a deviation from a normal distribution. We log transformed the data and ran the Shapiro-Wilk test again, which indicated no substantial deviation from a normal distribution. A Levene's test did not show any violation of homogeneity of variances ($F_{5,264} = 1.11$, $p = 0.35$).

A two-way repeated measures ANOVA revealed a significant effect of Content ($F_{2,88} = 10.7$, $p = 6.9E-05$, $partial\ \eta^2 = 0.20$). We also found a significant interaction between the effects of Event and Content ($F_{2,88} = 3.9$, $p = 2.4E-2$, $partial\ \eta^2 = 0.08$) on participant reaction times. Post-hoc analysis using pairwise t-test with Bonferroni corrections revealed that participants were significantly slower in reacting to physical events between conditions with Content: Blocks + Task and Content: Blocks ($t(44) = -2.88$, $p = 1.8E-2$, $Cohen's\ d = 0.43$). Participants were also significantly slower in reacting to physical events between conditions with Content: Blocks + Task and Content: None ($t(44) = -4.69$, $p = 7.9E-5$, $Cohen's\ d = 0.70$). We found no significant differences in reaction times between Event Type: Physical and Event Type: Virtual within all levels of Content (Figure 6).

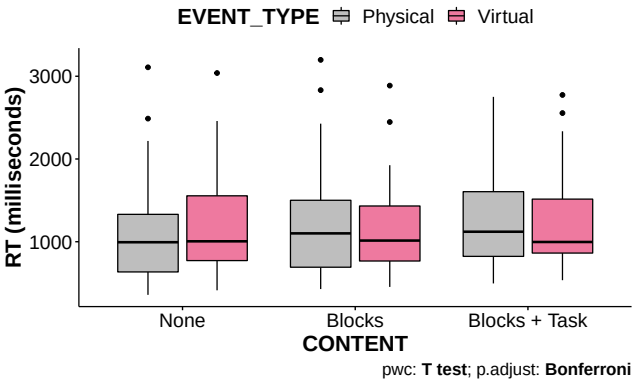

Figure 6: Boxplot of reaction times grouped by Content and Event Type for study 2. No significant differences in reaction times were found between Event Type: Physical and Event Type: Virtual for all levels of Content.

We conducted equivalence tests using the TOST procedure to determine if the effect of Event Type: Physical were statistically equivalent to the effect of Event Type: Virtual within the different levels of Content. The results of our equivalence test are detailed in Table 5.

Table 5: Results of the TOST equivalence test between conditions with Event Type: Physical and Event Type: Virtual, within each CONTENT level for study 2. Observed effects were statistically equivalent to zero for conditions Content: Blocks and Content: Blocks + Task between Event Type (Physical and Virtual). Observed effects were not statistically equivalent to zero for condition Content: None between Event Type (Physical and Virtual).

Content	Event 1	Event 2	t(88)	p-value
None	Physical	Virtual	1.117	$1.3E-1$
Blocks	Physical	Virtual	-2.006	$2.4E-2$
Blocks + Task	Physical	Virtual	-1.805	$3.7E-2$

Qualitative Results. We analysed the post-experiment questionnaire to explore factors that may have affected participant reaction times. We asked participants if they had any prior experience with AR applications to consider effects of familiarity with the technology. Only 13 out of the 48 participants mentioned that they briefly used AR applications, quoting Pokémon GO, Wizards Unite and Wanna Kicks, an application for trying out shoes. No participant reported extensive experience with AR technology.

We asked participants which conditions made it difficult for them to react to the LED. Similar to Study 1, participants unanimously agreed that Content: Blocks + Task was the most difficult. 4 out of 48 participants also mentioned that simply seeing the virtual blocks distracted them and may have affected their reaction to the events.

With respect to Event Type, 23 out of 48 participants reported that they did not notice any difference between the physical and

virtual LEDs. 19 out of 48 participants reported that it was easier to react to the virtual LEDs as they seemed brighter and the remaining 6 participants reported that it was easier to react to the physical LEDs.

5 DISCUSSION

The aim of this research was to understand how virtual content and the presence of an associated task influence the attentional tunneling effect on handheld AR applications. We operationalised attention in terms of reaction time [26], as the attentional tunneling phenomenon has been reported to cause increased reaction times to events.

To determine whether the attentional tunneling effect is caused by the presence of virtual content or the inclusion of an associated task, we measured reaction times to six different conditions based on Content and Event Type presented on a display. Our first study was administered through a handheld AR device to preserve ecological validity of the findings, and the second study was administered through a web-based application simulating an AR application to control for peripheral visibility of physical events.

Overall, we found that displaying virtual content in a handheld AR application does not significantly increase user reaction times to events, unless the virtual content presents information relevant to the user's task. We also found that peripheral visibility of physical events outside the frame of the handheld device leads to faster reaction to such events.

5.1 Effect of Content and Task on Attentional Tunneling

We analysed reaction time data for three different conditions based on Content: a baseline condition with no virtual content and no task (Content: None), a condition with virtual content and no task (Content: Blocks) and a condition with virtual content and a task (Content: Blocks + Task). In both of our studies, the task was to count the number of times a sphere was passed between blocks floating in virtual space. These conditions allowed us to test if the attentional tunneling effect is caused by the presence of virtual content by comparing reaction times between Content: None and Content: Blocks. It also allowed us to test if the attentional tunneling effect is caused by the presence of a task by comparing reaction times between Content: None and Content: Block + Task, and between Content: Blocks and Content: Block + Task.

The results from our first study indicate that there was a significant increase in reaction times when comparing conditions with a task (Content: Blocks + Task) to those without tasks (Content: None and Content: Blocks) for both physical and virtual Event Types. This suggests that attentional tunnelling occurs in the presence of a task associated to the virtual content in an AR application. This finding is aligned with prior work, which has shown that engagement with a task can cause attentional issues such as attentional tunneling [42] and inattentional blindness [46]. We also found that the observed effect between conditions with no additional task (Content: None and Content: Blocks) for both virtual and physical Event Types were statistically equivalent to zero. This supports our argument that the presence of virtual content alone does not cause the attentional tunneling effect. It also raises questions as to whether AR displays have solely caused the attentional tunneling effect as has been suggested in past work [16, 17].

The implications of these findings relate to how we evaluate AR applications and raise questions related to the kind of tasks that AR applications should be used to assist. Our findings indicate that the presence of a task within an AR application causes users to react slower (or miss) both virtual and physical events. This suggests that developers of handheld AR applications must consider the presence of tasks when attempting to alert users with virtual events or when expecting users to detect physical events. Our findings also suggest that virtual content can still be rendered on an AR application without significantly hindering event detection if the virtual content is not associated with a task. This means that designers of AR applications must consider the following questions in relation to the virtual content and the intended use of the AR applications.

First, does the virtual content present task-relevant information with respect to the intended use of the application? If the virtual content does present task-relevant information, it will draw users' attention and potentially reduce the likelihood of event detection or result in longer reaction times to events. However, content that is meant for aesthetic appeal (virtual graffiti on physical walls for example) and that is not associated with a task should minimally affect event detection.

The second question developers should ponder is: in what context is task-relevant virtual content displayed? Task-relevant virtual content should only be presented when it is ascertained that event detection is not crucial in terms of user safety and user experience. For example, designers of navigation applications using AR technology, such as Google Maps AR and Phiar [4], need to be particular mindful of user attention when evaluating their applications for user safety. Conversely, AR applications used in a relatively safer environment, such as IKEA Place [5], can afford to be less stringent with their evaluation in relation to user attention.

A curious observation in prior studies related to the attentional tunneling effect in different display technologies is that users excessively focus on the *virtual* content regardless of whether the users' task is related to the physical world view [16, 17] or within the virtual domain (our current work). We hypothesise that this is caused by the task-relevant information that the virtual content is used to present. In other words, users focus on the virtual contents of a display because the virtual content helps them complete their task. We plan to test this hypothesis in future work.

5.2 Effect of Event Type on Attentional Tunneling

We separately measured reaction times to physical and virtual events because prior work has reported that the attentional tunneling phenomenon increases reaction times to events presented on a channel that is different from the channel that caused the event [31, 51]. Specifically, if attentional tunneling was caused by virtual content in AR applications, we would expect to see shorter reaction times to events on the same channel (i.e., virtual events) when compared to physical events. However, results from both of our studies indicate that participants were not faster in reacting to virtual events when their reaction was supposedly meant to be 'tunneled' into the virtual content.

Our first study revealed that participants displayed significantly faster reaction times to physical events in comparison to virtual

events for all levels of Content. The results from our second study reveal that there is no statistically significant effect for conditions with virtual content (Content: Blocks and Content: Blocks + Task) between Event Type: Physical and Event Type: Virtual. These results suggest that the presence of virtual content in Content: Blocks and Content: Blocks + Task in both our studies does not significantly increase reaction times to virtual events as compared to physical events.

Interviews with participants from our first study highlight the role of peripheral visibility of the physical events to the faster reaction times. Participants reported that they could see the physical LEDs out of the corner of their eye even when they were visually focused on the screen of the smartphone, i.e., they could see the physical LED both on the smartphone screen and on their peripheral vision outside of the screen. Research on peripheral vision also highlights how peripheral vision is naturally used for motion and form changes of objects [49]. This may hint at the benefits of a smaller field of view of handheld AR devices at detecting physical events visible on user's peripheral view.

However, previous research has also suggested that the small field of view and VST nature of handheld smartphone devices may lead to increased cognitive load [11] therefore worsen the attentional tunneling effect [38, 52]. According to Baumeister et al., a conventional 2D monitor outperforms both VST and Optical See Through (OST) Head-Mounted Displays, while a Spatial Augmented Reality projection outperformed all other displays in a response time test. It is unclear, however, how a VST handheld AR device would fare against a VST AR HMD (e.g. Head mounted smartphone AR application) or against an OST AR HMD (e.g. HoloLens) in detecting physical events. While the VST/OST AR HMD provide a larger field of view and therefore should decrease cognitive load as compared to a handheld device, it could also be possible that the smaller field of view of a handheld device enables easier detection of physical events on account of peripheral vision.

On the note of peripheral vision, two participants in the first study were observed to rely solely on their peripheral vision to detect events. Their eye tracking recordings revealed that they would fixate their gaze at the center of the display during the no-task conditions and react to events without moving their gaze. They adopted a similar strategy for the task-based conditions where they would follow the sphere's movement with their gaze while reacting to events without switching their gaze to the event. The participants also stated their strategy in the interview while mentioning that it was harder to maintain with the virtually rendered events as they did not "stand out" as much in their peripheral vision. We suspect that this is because the physical LEDs lead to easier detection because they can be seen both within the smartphone's screen and outside of the display via peripheral vision. In contrast, the virtual LEDs were only visible within the device screen through peripheral vision. In our second study, two participants also reported using their peripheral vision to detect events, but mentioned that the LEDs were not easily detectable and thus they had to revert back to scanning for events. The significant increase in reaction times to physical events when they are visible outside the frame of the handheld device and the reports made by participants relating to peripheral vision highlight the impact peripheral visibility of events had on participants' reactions.

A participant in our first study also mentioned that the reflection of physical LEDs on a nearby surface made physical event detection easier than the virtual events. This property where an event can affect its surrounding can possibly be exploited to create a countermeasure for attentional tunneling via virtual content, i.e., graphical methods aimed at visualizing physical elements onto virtual objects may therefore help physical event detection. For example, visualizing physical environmental lighting is now possible on AR smartphones using APIs such as Google's ARCore [6].

5.3 Effect of Age, Modality and Supervision

One minor difference between our two studies was that participants in Study 2 exhibited larger mean reaction times to both physical and virtual events. A contributing factor to explain this difference is the higher average age (+8.1 years) of participants in Study 2, as compared to Study 1. Prior work has shown that age can lead to larger reaction times [23]. The larger reaction times may also be attributed to the change in input modality for reacting to events: an Arduino button in Study 1 versus a mouse in Study 2. Additionally, while the web-application used in Study 2 provided detailed instructions for valid task completion, the lack of face-to-face supervision may have also contributed to the larger reaction times in Study 2.

6 LIMITATIONS

Our studies presented participants with a task that required them to scan virtual content. As such, there are limitations in relation to the type of task we employ and the ability to generalise our findings to other display types.

First, the type of task employed in the study enabled us to test user reaction times in a controlled manner, and is reflective of tasks in AR applications that require users to scan the virtual content for task relevant information. However, the task was administered through a stationary display. This was necessary to tease apart the effects of task and content while controlling for the possible impact of mobility on attentional tunneling. Given that many AR applications are deployed in public settings, such as in the case of Pokémon GO [7] and Santa's Lil Helper [27], future studies should explore the prevalence of attentional tunneling in tasks where users have to move around their physical surroundings.

Second, our findings suggest that task causes the attentional tunneling effect and the attentional tunneling effect is not solely induced by the presence of virtual content in handheld AR VST devices, but these insights may not be generalizable to HUD and 3D displays. This is because virtual elements in AR are registered to the physical world [10] while virtual elements in HUD and 3D displays may not be and studies have already demonstrated how conformal imagery mitigates the attentional tunneling effect [31]. The finding, however, does question if further tests should be conducted to explore how task impacts the attentional tunneling effect in 3D displays and HUDs.

7 CONCLUSION

This paper reported two studies that aimed to tease apart the possible causes of the attentional tunneling effect in handheld AR applications. We found that the attentional tunneling effect is not caused by the mere presence of virtual content in an AR environment, but rather by the presence of tasks associated with the virtual content within the AR application. We also found that the small field of view of the smartphone device assisted users in detecting physical events

within their peripheral vision, i.e. outside of the smartphone's display. This can aid as a countermeasure for any attentional tunneling that may be caused by a task presented on the smartphone.

Our findings contribute towards better understanding of the attentional tunneling effect in handheld AR applications, and emphasise the need to consider tasks and physical context in an AR application to reduce user risks associated with the attentional tunneling effect. In essence, careful considerations must be made when rendering virtual content with task relevant information in situations where event detection is essential for user safety and user experience.

ACKNOWLEDGEMENTS

We thank our participants for their time and the reviewers for their feedback that helped us improve our paper. This research was supported by the Melbourne Research Scholarship provided by the University of Melbourne. Eduardo Velloso is the recipient of an Australian Research Council Discovery Early Career Award (Project Number: DE180100315) funded by the Australian Government.

REFERENCES

[1] [n.d.]. Arduino. https://www.arduino.cc/.
[2] [n.d.]. Consumer mobile device augmented reality applications (embedded/standalone). https://www.statista.com/statistics/608967/mobile-ar-applications-installed-base-worldwide/.
[3] [n.d.]. Highway in the Sky. https://www.hq.nasa.gov/office/aero/docs/chicago/hits.htm.
[4] [n.d.]. Phiar. https://www.phiar.net/.
[5] 2014. IKEA Place. https://www.ikea.com/au/en/apps/IKEAPlace.html.
[6] 2019. ARCore, Google. https://developers.google.com/ar/.
[7] Niantic, Inc. 2016. Pokémon GO. https://www.pokemongo.com.
[8] Paul Atchley and Jeff Dressel. 2004. Conversation limits the functional field of view. *Human factors* 46, 4 (2004), 664–673.
[9] John W Ayers, Eric C Leas, Mark Dredze, Jon-Patrick Allem, Jurek G Grabowski, and Linda Hill. 2016. Pokémon GO—a new distraction for drivers and pedestrians. *JAMA internal medicine* 176, 12 (2016), 1865–1866.
[10] Ronald T Azuma. 1997. A survey of augmented reality. *Presence: Teleoperators & Virtual Environments* 6, 4 (1997), 355–385.
[11] James Baumeister, Seung Youb Ssin, Neven AM ElSayed, Jillian Dorrian, David P Webb, James A Walsh, Timothy M Simon, Andrew Irlitti, Ross T Smith, Mark Kohler, et al. 2017. Cognitive cost of using augmented reality displays. *IEEE transactions on visualization and computer graphics* 23, 11 (2017), 2378–2388.
[12] Mark Billinghurst and Hirokazu Kato. 2002. Collaborative augmented reality. *Commun. ACM* 45, 7 (2002), 64–70.
[13] Salvador Bueno, M Dolores Gallego, and Jan Noyes. 2020. Uses and Gratifications on Augmented Reality Games: An Examination of Pokémon Go. *Applied Sciences* 10, 5 (2020), 1644.
[14] George Chang, Patricia Morreale, and Padmavathi Medicherla. 2010. Applications of augmented reality systems in education. In *Society for Information Technology & Teacher Education International Conference*. Association for the Advancement of Computing in Education (AACE), 1380–1385.
[15] Kuo-En Chang, Chia-Tzu Chang, Huei-Tse Hou, Yao-Ting Sung, Huei-Lin Chao, and Cheng-Ming Lee. 2014. Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum. *Computers & Education* 71 (2014), 185–197. https://doi.org/10.1016/j.compedu.2013.09.022
[16] Benjamin J Dixon, Michael J Daly, Harley Chan, Allan D Vescan, Ian J Witterick, and Jonathan C Irish. 2013. Surgeons blinded by enhanced navigation: the effect of augmented reality on attention. *Surgical endoscopy* 27, 2 (2013), 454–461.
[17] Benjamin J Dixon, Michael J Daly, Harley HL Chan, Allan Vescan, Ian J Witterick, and Jonathan C Irish. 2014. Inattentional blindness increased with augmented reality surgical navigation. *American journal of rhinology & allergy* 28, 5 (2014), 433–437.
[18] Matt Dunleavy, Chris Dede, and Rebecca Mitchell. 2009. Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of science Education and Technology* 18, 1 (2009), 7–22. https://doi.org/10.1007/s10956-008-9119-1
[19] Erik Edström, Gustav Burström, Artur Omar, Rami Nachabe, Michael Söderman, Oscar Persson, Paul Gerdhem, and Adrian Elmi-Terander. 2020. Augmented reality surgical navigation in spine surgery to minimize staff radiation exposure. *Spine* 45, 1 (2020), E45–E53.
[20] Lizbeth Escobedo, Monica Tentori, Eduardo Quintana, Jesus Favela, and Daniel Garcia-Rosas. 2014. Using augmented reality to help children with autism stay focused. *IEEE Pervasive Computing* 13, 1 (2014), 38–46.
[21] Steven Fadden, Patricia May Ververs, and Christopher D. Wickens. 2001. Pathway HUDs: Are They Viable? *Human Factors* 43, 2 (2001), 173–193. https://doi.org/10.1518/001872001775900841 arXiv:https://doi.org/10.1518/001872001775900841 PMID: 11592660.
[22] Edith Fischer and Richard F Haines. 1980. Cognitive issues in head-up displays. (1980).
[23] James L Fozard, Max Vercruyssen, Sara L Reynolds, PA Hancock, and Reginald E Quilter. 1994. Age differences and changes in reaction time: the Baltimore Longitudinal Study of Aging. *Journal of gerontology* 49, 4 (1994), P179–P189.
[24] Henry Fuchs, Mark A Livingston, Ramesh Raskar, Kurtis Keller, Jessica R Crawford, Paul Rademacher, Samuel H Drake, Anthony A Meyer, et al. 1998. Augmented reality visualization for laparoscopic surgery. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 934–943.
[25] Kristiina Jokinen and Topi Hurtig. 2006. User expectations and real experience on a multimodal interactive system. In *Ninth International Conference on Spoken Language Processing*.
[26] Stephanie AH Jones, Beverly C Butler, Franziska Kintzel, Anne Johnson, Raymond M Klein, and Gail A Eskes. 2016. Measuring the performance of attention networks with the Dalhousie Computerized Attention Battery (DalCAB): Methodology and reliability in healthy adults. *Frontiers in psychology* 7 (2016), 823.
[27] Ryan M. Kelly, Hasan Shahid Ferdous, Niels Wouters, and Frank Vetere. 2019. Can Mobile Augmented Reality Stimulate a Honeypot Effect? Observations from Santa's Lil Helper. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. ACM, New York, NY, USA, Article 285, 13 pages. https://doi.org/10.1145/3290605.3300515
[28] Elisa Maria Klose, Nils Adrian Mack, Jens Hegenberg, and Ludger Schmidt. 2019. Text presentation for augmented reality applications in dual-task situations. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 636–644.
[29] Sean W Kortschot and Greg A Jamieson. 2019. Classification of Attentional Tunneling Through Behavioral Indices. *Human factors* (2019), 0018720819857266.
[30] JW Lasswell and CD Wickens. 1995. The effects of display location and dimensionality on taxi-way navigation (Tech. Rep. No. ARL-95-5/NASA-95-2). *Savoy: University of Illinois, Aviation Research Laboratory* (1995).
[31] Jonathan L Levy, David C Foyle, and Robert S McCann. 1998. Performance benefits with scene-linked HUD symbology: an attentional phenomenon?. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 42. SAGE Publications Sage CA: Los Angeles, CA, 11–15.
[32] Ana Paz Goncalves Martins, Steffen Hölscher, and Thomas Dautermann. 2020. Evaluation of a Tunnel-in-the-Sky Head-Up Display Design for Curved Approaches Using Eye-Tracking. *Aviation Psychology and Applied Human Factors* (2020).
[33] Kenneth C Mills, Susan E Spruill, Roy W Kanne, Katherine M Parkman, and Ying Zhang. 2001. The influence of stimulants, sedatives, and fatigue on tunnel vision: risk factors for driving and piloting. *Human factors* 43, 2 (2001), 310–327.
[34] Oscar Olmos, Christopher D Wickens, and Andrew Chudy. 2000. Tactical displays for combat awareness: An examination of dimensionality and frame of reference concepts and the application of cognitive engineering. *The International Journal of Aviation Psychology* 10, 3 (2000), 247–271.
[35] Thomas Olsson, Else Lagerstam, Tuula Kärkkäinen, and Kaisa Väänänen-Vainio-Mattila. 2013. Expected user experience of mobile augmented reality services: a user study in the context of shopping centres. *Personal and Ubiquitous Computing* 17, 2 (2013), 287–304. https://doi.org/10.1007/s00779-011-0494-x
[36] Lyn Pemberton and Marcus Winter. 2009. Collaborative augmented reality in schools. In *Proceedings of the 9th international conference on Computer supported collaborative learning-Volume 2*. International Society of the Learning Sciences, 109–111.
[37] LJ Prinzel and LJ Kramer. 2009. Synthetic vision systems. (2009).
[38] Esa M Rantanen and Joseph H Goldberg. 1999. The effect of mental workload on the visual field size and shape. *Ergonomics* 42, 6 (1999), 816–834.
[39] Stuart Reeves, Mike Fraser, Holger Schnadelbach, and Steve Benford. 2005. Engaging Augmented Reality in Public Places. In *Adjunct Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, Oregon, USA) *(CHI '05)*. ACM, New York, NY, USA.
[40] Nicolas Régis, Frédéric Dehais, Emmanuel Rachelson, Charles Thooris, Sergio Pizziol, Mickaël Causse, and Catherine Tessier. 2014. Formal detection of attentional tunneling in human operator–automation interactions. *IEEE Transactions on Human-Machine Systems* 44, 3 (2014), 326–336.
[41] Joceline Roge, Laetitia Kielbasa, and Alain Muzet. 2002. Deformation of the useful visual field with state of vigilance, task priority, and central task complexity. *Perceptual and motor skills* 95, 1 (2002), 118–130.
[42] Sylvia Rothe, Daniel Buschek, and Heinrich Hußmann. 2019. Guidance in Cinematic Virtual Reality-Taxonomy, Research Status and Challenges. *Multimodal Technologies and Interaction* 3, 1 (2019), 19. https://doi.org/10.3390/mti3010019
[43] Julie Saint Lot, Jean-Paul Imbert, and Frédéric Dehais. 2020. Red Altert: a cognitive countermeasure to mitigate attentional tunneling. In *Proceedings CHI*. 25–30.

[44] Donald J Schuirmann. 1987. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of pharmacokinetics and biopharmaceutics* 15, 6 (1987), 657–680.

[45] Michael A Seaman and Ronald C Serlin. 1998. Equivalence confidence intervals for two-group comparisons of means. *Psychological methods* 3, 4 (1998), 403.

[46] Daniel J Simons and Christopher F Chabris. 1999. Gorillas in our midst: Sustained inattentional blindness for dynamic events. *perception* 28, 9 (1999), 1059–1074.

[47] Brandon Victor Syiem, Ryan M Kelly, Eduardo Velloso, Jorge Goncalves, and Tilman Dingler. 2020. Enhancing Visitor Experience or Hindering Docent Roles: Attentional Issues in Augmented Reality Supported Installations. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 279–288. https://doi.org/10.1109/ISMAR50242.2020.00053

[48] Marcus Tonnis, Christian Sandor, Christian Lange, and Heiner Bubb. 2005. Experimental Evaluation of an Augmented Reality Visualization for Directing a Car Driver's Attention. In *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '05)*. IEEE Computer Society, USA, 56–59. https://doi.org/10.1109/ISMAR.2005.31

[49] Christian Vater, Ralf Kredel, and Ernst-Joachim Hossner. 2016. Detecting single-target changes in multiple object tracking: The case of peripheral vision. *Attention, Perception, & Psychophysics* 78, 4 (2016), 1004–1019.

[50] Victoria R Wagner-Greene, Amy J Wotring, Thomas Castor, Jessica Kruger, Sarah Mortemore, and Joseph A Dake. 2017. Pokémon GO: Healthy or harmful? *American journal of public health* 107, 1 (2017), 35.

[51] Christopher D Wickens and Amy L Alexander. 2009. Attentional tunneling and task management in synthetic vision displays. *The International Journal of Aviation Psychology* 19, 2 (2009), 182–199. https://doi.org/10.1080/10508410902766549

[52] LJ Williams. 1995. Visual field narrowing induced by workload. *J Gen Psychol* 122 (1995), 225–235.

[53] Michelle Yeh, James L Merlo, Christopher D Wickens, and David L Brandenburg. 2003. Head up versus head down: The costs of imprecision, unreliability, and visual clutter on cue effectiveness for display signaling. *Human Factors* 45, 3 (2003), 390–407.

[54] Michelle Yeh and Christopher D Wickens. 2001. Display signaling in augmented reality: Effects of cue reliability and image realism on attention allocation and trust calibration. *Human Factors* 43, 3 (2001), 355–365.

[55] Qiushi Zhou, Difeng Yu, Martin N Reinoso, Joshua Newn, Jorge Goncalves, and Eduardo Velloso. 2020. Eyes-free Target Acquisition During Walking in Immersive Mixed Reality. *IEEE Transactions on Visualization and Computer Graphics* 26, 12 (2020), 3423–3433.