Virtual Field Studies: Conducting Studies on Public Displays in Virtual Reality

Ville Mäkelä[1,5], Rivu Radiah[2], Saleh Alsherif[3], Mohamed Khamis[4], Chong Xiao[1], Lisa Borchert[1], Albrecht Schmidt[1], Florian Alt[2]

[1]LMU Munich, {ville.maekelae, albrecht.schmidt}@ifi.lmu.de, {chong.xiao, lisa.borchert}@campus.lmu.de
[2]Bundeswehr University Munich, {sheikh.rivu, florian.alt}@unibw.de
[3]German University in Cairo, saleh.alsheeref@student.guc.edu.eg
[4]University of Glasgow, mohamed.khamis@glasgow.ac.uk
[5]Tampere University

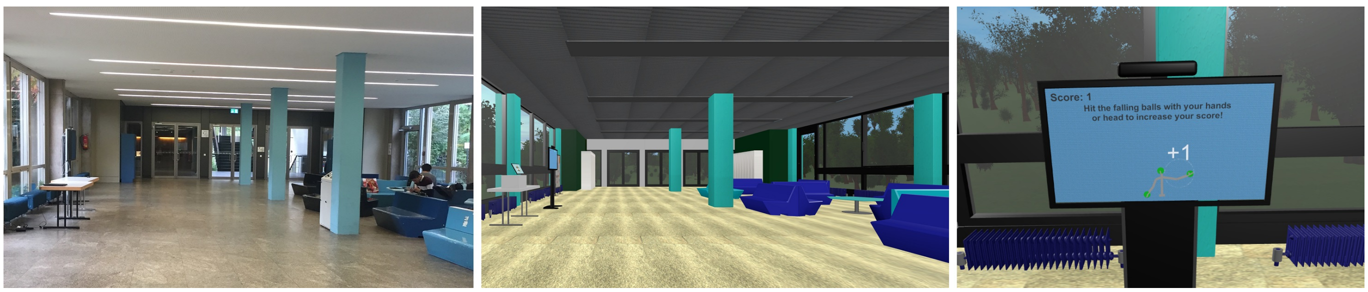

Figure 1. We explore whether field studies on public displays can be conducted in virtual reality. In two user studies we compare user behavior between a real public space (left) and a virtual public space (middle). For one study, we developed a gesture-controlled display for both environments (right).

ABSTRACT
Field studies on public displays can be difficult, expensive, and time-consuming. We investigate the feasibility of using virtual reality (VR) as a test-bed to evaluate deployments of public displays. Specifically, we investigate whether results from virtual field studies, conducted in a virtual public space, would match the results from a corresponding real-world setting. We report on two empirical user studies where we compared audience behavior around a virtual public display in the virtual world to audience behavior around a real public display. We found that virtual field studies can be a powerful research tool, as in both studies we observed largely similar behavior between the settings. We discuss the opportunities, challenges, and limitations of using virtual reality to conduct field studies, and provide lessons learned from our work that can help researchers decide whether to employ VR in their research and what factors to account for if doing so.

Author Keywords
Virtual reality; field studies; public displays; research methods.

CCS Concepts
•**Human-centered computing** → **User studies; VR;**

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

ACM ISBN 978-1-4503-6708-0/20/04 ...$15.00.
http://dx.doi.org/10.1145/3313831.3376796

INTRODUCTION
Public displays have received considerable attention from the research community in the past years and are being increasingly deployed in public spaces [21]. While there is today a general understanding of the many challenges that evolve around public displays, much of the gained knowledge in our research community is still difficult to generalize as data in field studies are usually gathered in one specific setting. As a result, audiences are widely different (ranging from children, via adults to older adults), displays are encountered in different situations (for example, a waiting situation at a bus stop vs. people being in a hurry in a busy street), are frequented by different audiences at different day times, and the audience may be affected by many other things that happen nearby (a street festival, a construction site, vehicles parked in front of the display, etc.). While all these situations yield interesting insights, it is often desirable to have more control to be able to investigate the influence of different contexts in more detail.

At the same time, public display deployments pose considerable effort to researchers [29, 54, 72]: researchers face the challenge of finding suitable locations for their deployment [24,32], as a result of which deployments are often opportunistic [11,26]; hardware needs to be maintained and software run in a stable manner [43]; and data collection is time-consuming as researchers need to be present in the display vicinity for observations and subsequent interviews [8,43].

A new challenge arises for researchers as a result of new privacy protection regulations, such as the European GDPR [5]. These regulations require informed consent to be obtained from people prior to data being collected about them. This introduces new legal and ethical considerations [75], which in turn may significantly complicate study processes in public spaces and may make it difficult to obtain relevant permissions.

In this paper, we investigate Virtual Reality (VR) as a research paradigm for conducting public display studies. VR allows for a high degree of control over the environment, i.e., researchers can easily manipulate the physical layout of a space, the number of people in the vicinity, and physical properties of the display (size, orientation, resolution), etc. This led us to believe that traditional approaches to public display research can be complemented by an investigation in VR.

The particular focus of our work is on understanding the strengths and weaknesses of this approach: for which types of common research questions [9] is the approach viable and for which is it not? How ecologically valid is the data? How is the experience of users different in real and virtual environments? We believe this understanding to be valuable for the HCI community in general and for the public display community in particular, since it enables an entirely new approach to research, the understanding of which is a crucial initial step.

In this work, we first provide an *in-depth introduction to the concept*, specifically discussing the potential of VR for different types of research question related to public displays. Then, we present *two comparative studies*, designed to shed light on different aspects of this new research paradigm. In both studies, we achieve this by investigating the participants' behavior under similar conditions in a real-world public space and a virtual public space. The first study (N=40) is focused on investigating audience behavior. In particular, we investigate whether the well-known honeypot effect [16,49] can be recreated in VR and whether similar behavior can be observed. The second study (N=16) investigates display effectiveness [9], for example, if and how users notice, approach, and interact with a display in a waiting situation.

The results of our early investigation suggest that *virtual field studies can be a powerful research tool*, as the results can transition over to the real world. In Study 1, we confirmed that the honeypot effect can occur in VR. In Study 2, we observed that users noticed, approached, and interacted with a virtual public display in a largely similar manner as they did in the real world. A notable difference was found in that users are more motivated to explore virtual environments, and may be more likely to interact with a public display in VR.

Based on our work, we provide *three lessons learned* regarding virtual field studies, aimed at helping researchers and practitioners to 1) decide whether or not to employ VR in their studies, 2) design study tasks that stimulate natural behavior in virtual reality, and 3) interpret results from VR studies.

CONTRIBUTION STATEMENT

The contribution of our work is threefold: (1) We introduce the concept of employing VR as a research paradigm for public display studies. (2) We present two studies, comparing an investigation in the real world to the same investigation in VR, to understand both strengths and pitfalls. (3) We distill lessons learned to help public display researchers design, conduct and interpret results from further studies in VR.

RELATED WORK

Our work draws from previous work on 1) evaluation of public displays, 2) audience behavior around displays, and 3) the use of VR as a study platform.

Evaluation of Public Displays

Different study paradigms exist to evaluate public displays. A summary is provided by Alt et al. [9]. The majority of public display research employs lab studies [14,40,51,58] and field studies [28,52,64–66,73,76,82]. Generally, the paradigm is chosen based on the research question. Lab studies are suitable for conducting evaluations in controlled environments. Examples include evaluation of novel interaction mechanisms [1,53] or studying user behavior when presented with certain content [6,37]. Lab studies allow controlled collection of data but suffer from low ecological validity. On the other hand, field studies are conducted in public spaces and offer more ecologically valid results. Unlike lab studies, field studies are suitable for studying natural behavior [2,3,26,49].

For example, Dalton et al. [20] studied if passersby notice displays by equipping participants with eye trackers and asking them to walk through a building with multiple public displays. The downside of field studies is that they are costly [21,23,34, 52,59], time consuming [9], hard to anticipate (e.g., behavior of passersby) [26], and subject to a plethora of external factors that may negatively impact the display and the validity of the collected results [54]. Field studies often require medium to long-term deployments, hence requiring robust hardware and software that can operate without attendance for several weeks or months. Multiple works reported on challenges of (long-term) deployments [33,43,54,72].

In this work, we study the use of VR as a test bed to evaluate public displays. Like lab studies, VR studies can be controlled, which allows for accurate collection of data and prevents many pitfalls of real-world public deployments. At the same time, VR studies bring some of the benefits of field studies in terms of scale of the studies, and the ability to study natural behavior. Unlike lab and field studies, VR allows easy manipulation of the setup as well as the behavior of passersby. This work aims to understand how well outcomes from VR studies match those collected in real world deployments.

Audience Behavior around Public Displays

The behavior of users and bystanders around public displays has been studied extensively. Most notably, researchers explored how 1) social aspects, 2) environmental variables, and 3) the displayed content impact on audience behavior.

One of the most studied social behavior around public displays is the *honeypot effect* [15,16,26,41,49,83], which occurs when users engaging with the display implicitly attract the attention of others in the vicinity of the display [16]. Another example is the *staging effect*; some might refrain from interacting with a display due to social embarrassment [18], while others might interact only to be seen by peers [19]. Audience behavior is also influenced by environmental factors, such as the arrangement of the space. For example, Gentile et al. [26] studied a setup where a public display was surrounded by seats. They found that users position themselves farther away from the display when said seats are occupied. Dalton et al. [20] found that the user's awareness of the display is influenced by the architecture of the building in which it is deployed. The way displays are configured and placed influence how users position themselves before interacting with them [59, 74], and could even entice users to communicate even if they do not know each other [60, 79]. Müller et al. [49] found placing displays near certain elements impact attention to the display. To combat *display blindness* [42], previous work manipulated the display's content to attract attention to the display. Multiple works used mirrored user representations to attract the passerby's attention [49, 81]. Others used flashing objects [36], subtle gaze direction techniques [71], and moving physical objects [35] to attract attention to the display.

Out of all the aforementioned aspects that impact audience behavior, the honeypot effect is by far the most studied one. This motivated us to investigate if the effect occurs in virtual reality as well. Thus, in Study 1, we studied the user's response to virtual avatars surrounding displays in a virtual environment.

VR as a Study Platform

Prior work has investigated the use of VR as a platform to acquire results, knowledge, or skills that transfer to the real world either partially or completely.

Some prior research has aimed to compare how users behave in VR as opposed to comparable real-world scenarios. Moussaid et al. [45] used VR to understand crowd behavior during a high-stress evacuation scenario. Their results show similar crowd behavior in high-risk situations in VR compared to real situations. Schrom-Feiertag et al. [68, 69] studied how immersive virtual environments can be leveraged to evaluate public infrastructures using a mobile eye tracking system. They found that participants in real and virtual train stations exhibit similar wayfinding behavior in terms of decision making and attention. Deb et al. [22] used VR to simulate a pedestrian crossing. Their collected objective measures, such as walking speed, were similar to real-world norms. Agethen et al. [4] conducted similar research where they studied how immersion in a virtual environment affects human locomotion. A comparison with the real world concluded that VR can be used as an evaluation tool for analyzing human locomotion.

Virtual reality has also been studied for education purposes [27, 57]. For example, Gorecky et al. [27] found that improving performance of factory workers during VR training results in better performance in similar real-world tasks. Other uses for VR have been the evaluation of the intrusiveness of an advertising app using a virtual supermarket [34], and the use of virtual assistance to evaluate pervasive applications [12].

These studies bring forth the potential of using VR as a research platform. Most importantly, prior work shows that users can adopt similar behaviors in VR as they do in the real world in a variety of scenarios, which in turn suggests that employing VR to evaluate real-world deployments could be feasible. However, to date, the use of VR to evaluate public displays has not been explored. This underlines the importance of understanding how transferable findings from VR studies are to real-world deployments of public displays.

VIRTUAL PUBLIC DISPLAY STUDIES

Much work exists on the evaluation of aspects that relate to public displays and user behavior around them. For an in-depth introduction to methodology, we refer to Alt et al. [9] and Davies et al. [21]. We contribute to the public display research methodology by investigating how (a) virtual reality compares to existing research paradigms, such as lab studies and field studies and (b) for which type of the research questions is VR a suitable complement or even replacement.

Alt et al. [9] distinguish research questions on audience behavior, user experience, user acceptance, user performance, display effectiveness, privacy, and social impact. In the following, we provide a brief discussion on the potential strengths and weaknesses of VR regarding these questions. This reflection ultimately informed the design and focus of our work.

Audience Behavior. Much work in recent years has focused on the behavior of people in the vicinity of interactive displays, discovering that phenomena such as the honeypot effect [16], the staging effect [19] or the butt brush effect [78] exist also for public displays, but also finding new effects, such as the landing effect [49]. Most of this work was conducted in public spaces where people behaved naturally. This raises the question how to support such natural behavior in Virtual Reality. A particular challenge here is to not make it obvious to people that the public display is the object under investigation. At the same time, other factors play a role, including other people in the space, other objects striving for the attention of users, and the ability to move freely.

User Experience. User experience is often assessed as a side question in public display research. In particular, researchers often ask that participants fill in standardized questionnaires such as AttrakDiff [31] or UEQ [67] after people have experienced a particular aspect related to public displays. On one hand, this might require thinking about how filling such questionnaires can be embedded in VR in a suitable manner [70]. At the same time, a challenge might be that perceiving public displays in VR might in itself have an influence on the experience. Here it is important to think about phrasing questions in a way such that a clear distinction between the VR experience and the display experience becomes possible.

User Acceptance. To understand whether users would accept a certain technology related to public displays, researchers often conduct focus groups. We believe virtual reality to be a powerful tool here, since a more realistic presentation can be achieved compared to a lab setting.

User Performance. When it comes to measuring performance, prior work has usually favored studies in the lab over studies in the field, due to the ability to better control confounding factors [7, 10]. Here, we expect one of the major strengths of VR, since it allows for fine-grained control over the environment, such as number of people, their trajectories, their behavior, etc. while still providing a more realistic experience compared to a lab setting (e.g., simulating crowded areas). Also, VR can allow quick and low-effort testing of many different conditions (e.g., different display sizes, orientations, different degree of crowdedness) and hence obtain more generalizable results.

Display Effectiveness. Assessing display effectiveness has been at the focus of researchers since the advent of interactive displays. Almost all deployments report on the number of people seeing the display, approaching the display, and interacting with the display. Researchers also propose standardized ways of reporting display effectiveness [42]. We believe virtual reality to be well suited for assessing effectiveness, because gaze direction, distance to the display and engagement/interaction can be easily logged, compared to prior work where this is often done manually by inspecting recorded videos.

Privacy. This aspect received relatively little attention in the past. While some approaches to protect users' privacy in front of public displays have been proposed, investigation has usually been done via demonstrations in the lab. This might be a result of the fact that privacy is difficult to study in the field without unethically putting users (and their data) at risk. With VR, solutions can be presented in a more realistic manner and users' privacy perception be considered more reliably.

Social Impact. Prior research has looked into which social behavior evolves around public displays. This includes both cases where a public display application was intentionally designed for fostering social engagement and cases where this was unexpected. We expect the investigation of related questions in VR to be feasible, yet challenging, since multiple users would need to be present (e.g., to interact with each other). Wearing a head-mounted display (HMD) and not being able to see others' reaction might be a major challenge. Creating realistic virtual agents is challenging [25].

Research Approach

The previous section demonstrates that there are many opportunities but also challenges in answering research questions related to public displays in virtual reality. To begin addressing this broad research area, in this paper we focus on *audience behavior* and *display effectiveness*.

We chose these aspects for two main reasons. First, they are most commonly addressed in field studies [9]; complementing or even replacing field studies is where the benefit of virtual reality studies is most apparent. Second, these aspects deal with fundamental behavioral questions, for example, whether and how people notice, approach, or interact with a public display. These aspects relate to the various phases of interaction, cf. the Audience Funnel framework [44]. Understanding such basic behaviors is imperative to understanding behavior in virtual reality more generally – in particular, how it matches and how it differs from real-world behavior.

To this end, we designed two studies where we put participants under similar conditions in a real-world public space and a virtual public space, to investigate the similarities and differences between the two. In the first study, we focused on audience behavior and the early steps of interaction (e.g., attention) [44]. In particular, we created a virtual honeypot effect and investigated whether and how this influences participants' behavior and attention towards the display. In the second study, we focused on display effectiveness and overall behavioral analysis, for example, how users approach and interact with the display, and how they transition from one interactive phase to another.

Figure 2. Actors standing in front of a public display in the real-world condition of Study 1 to stimulate the honeypot effect.

STUDY 1: THE HONEYPOT EFFECT IN VR

The objective of the first study was to compare audience behavior in the real world and in VR. In particular, we investigated the existence and nature of the honeypot effect. We chose the honeypot effect due to being a widely known and well-explored phenomenon that has previously received considerable attention from the public display community [16, 28, 36, 38, 47–49, 56, 73, 83].

Study Design

We recruited 40 participants (21 females) with an average age of 22.8 (SD=4.0) for our user study using university mailing lists. The study was designed as a between-subjects experiment with two independent variables: the environment (real world vs. VR) and the existence of the honeypot effect (yes / no). Twenty of the participants experienced the real world, and the other twenty experienced the virtual world. Each group of 20 was further split to 10 who experienced the honeypot effect, and the other 10 did not experience a honeypot effect. We used actors as a live audience in the real world (Figure 2) and virtual characters as a digital audience in VR (Figure 3) to create the honeypot conditions. We followed a between-subjects design because the audience appearing or leaving mid-study might have revealed them as part of the study. Participants were compensated with a 5€ online shop voucher.

As dependent variables, we measured the attention towards the public display in two ways. First, we collected gaze data from participants. For VR users, this was done using a Pupil Labs eye tracker integrated in the HTC Vive HMD. For real-world users, we used the Tobii Pro Glasses. Furthermore, we asked participants to describe the contents on the public display after they had finished their study tasks, to evaluate whether they had focused on what was showing on the display.

Apparatus

The real-world condition took place at the lobby of a university building in a popular urban area (Figure 1, left). The lobby is used as a through-pass between class rooms, computer rooms, and the street outside. The same space is also equipped with tables, chairs, and snack and drink machines, and is often used by students to study and socialize. As such, the location is a good representation of spaces commonly used in public display research [26, 38, 47, 48, 50, 52, 73, 79, 80].

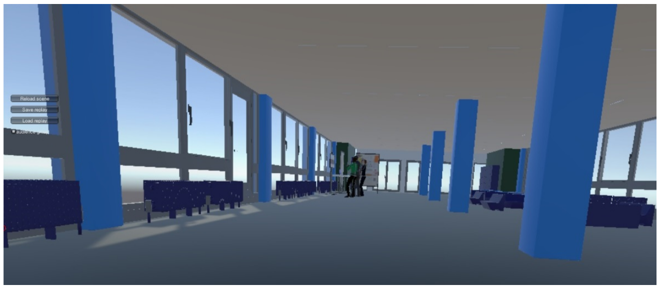

Figure 3. The virtual location in Study 1 in the honeypot effect condition. Participants had to walk past the display, in front of which animated virtual characters were standing and looking at it.

In the real-world condition, participants were equipped with the Tobii Pro Glasses 2 wearable eye tracker to record their gaze data during the experiment. In the VR condition, we used the HTC Vive as the head-mounted display and its built-in eye tracker to record the participants' gaze. Participants in the VR group conducted the study in a lab.

For the VR condition, we created a virtual replica of the space, using the Unity platform (see Figure 3). We focused on replicating the layout (size and shape) and important objects in the space (pillars, windows, doors, tables, chairs, lockers, and snack machines). The virtual environment was also populated with animated virtual characters walking around the space to match the dynamics of the real world condition. The study was conducted during a relatively quiet time between semesters where only few bystanders were present in the real world condition and distraction was minimal.

The VR condition participants needed to be able to safely move through a large virtual environment. Participants used the Vive controller's touchpad to move left, right, forward and backwards. Orientation and viewport movement was controlled normally by moving and turning the head. This control scheme was decided based on a pilot study where 20 participants experienced: a) normal walking followed by redirection at boundaries [62], b) walking in place [77], c) large skips in VR in response to normal walking, and d) using the controller while stationary. The last was found the most realistic, comfortable and natural according to responses to Likert-scale questions. It was also the least demanding according to responses to a NASA TLX questionnaire [30].

Task

To ensure participants remained unaware of the public display being the subject of the investigation, their task was to find a word written on a note placed behind a pillar, return to the starting point, and report the word to the experimenter. The display was midway between the starting point and the pillar, which means participants passed by the display on their way to the pillar and on their way back. To collect more data, the task was repeated twice. The note as well as the content shown on the public display was changed each time. In the first trial, the public display showed a technology article, and an article about a sports match in the second one. The display was not mentioned at any point. The audiences (actors and virtual avatars) were positioned close enough to the display so that they did not affect the participants' walking path.

Procedure

Participants were first briefed about the study and asked to sign a consent form. For VR users, we gave a brief introduction to the VR headset and explained how to navigate. We then explained the task to the participants and that it would be completed twice. Next, the eye tracker was calibrated. For half of the participants, no one was present in front of the display to simulate the *no honeypot effect* condition. For the other half (*honeypot effect* condition), three actors stood in front of the public display and looked at it in the real world condition (Figure 2). In VR, we implemented the same condition by adding three animated virtual characters who behaved similarly to their real-world counterparts (Figure 3). To ensure similarity, the real-world actors and the virtual agents did not speak, and they did not interact with the study participants in any way.

After finishing the tasks, participants filled in a questionnaire and we conducted a semi-structured interview in which we revealed our interest in the display for the first time.

Results

Attention Towards the Public Display

Using the collected gaze data, we measured how long participants gazed towards the public display. In the *real-world environment*, without an audience, participants looked towards the display on average for 1.23 seconds (SD=1.75 s), both search tasks included. With an audience, the same number was 1.46 seconds (SD=1.7 s) – an 18.7% increase in duration. In the *virtual environment*, the duration was 4.61 seconds (SD=5.7 s) without an audience, and 10.61 seconds (SD=8.7 s) with an audience — a 230.15% increase in duration.

A two-way ANOVA shows a statistically significant main effect of honeypot effect $F_{1,34} = 4.32, p < 0.001, \eta_p^2 = 0.113$ and of the environment$F_{1,34} = 18.41, p < 0.001, \eta_p^2 = 0.351$ on how long participants gazed at the display. There was no significant interaction between both factors ($p > 0.05$). Pairwise comparisons with Bonferroni correction indicate that in both environments gaze durations are significantly longer in the presence of an audience than in their absence ($p < 0.05$). Gaze durations are also significantly longer in VR compared to the real-world ($p < 0.001$).

Comprehension of Display Content

After the tasks, participants were asked to describe the display content. Any factual statement about the content was counted as a success (e.g., "it was about sports"). In the *real-world environment*, without an audience, 4 participants (40%) could successfully describe the content, while the other 6 could not. With an audience, 8 participants (80%) could do the same. In *VR*, only one participant (10%) comprehended the display's content when no audience was present. With an audience, 5 participants (50%) had looked at and understood the content.

Assessment of Audience Effects

We also asked participants how they thought an audience affected their perception of the public display, or in the case where no audience was present, how they thought an audience would affect their perception of the display. In all four groups, nine out of ten participants (90%) reported that an audience made – or would have made – the display more interesting.

Discussion and Summary

In both conditions, we found an increase in attention towards the public display as well as an increase in the understanding of display content when an audience was present. Moreover, participants reported similarly in both conditions that an audience did make the display more attractive. Therefore, *the honeypot effect was found in the VR condition*.

Still, we observed two differences in the results between the real and the virtual environment, which we discuss below:

The overall attention times were much higher in VR and the honeypot effect was stronger. This may have been influenced by three factors. First, movement in VR using a controller was slower compared to normal walking and, therefore, the time during which participants could observe the display was longer. Second, users might overall be more motivated to observe their surroundings in virtual reality compared to real-world environments. This likely made them interested in observing the virtual audience and, consequently, the display. Third, the low fidelity of the environment (e.g., no textures in most of the objects) may have driven participants' attention more towards the display.

Another difference was that despite longer average attention towards the display, fewer participants were able to comprehend the display's content in VR (10% without audience and and 50% with audience), compared to the real environment (40% and 80%, respectively). This may be due to the limitations of the VR hardware, as the lower resolution makes it more difficult to see distant display content clearly.

STUDY 2: BEHAVIOR AROUND A VR PUBLIC DISPLAY

In the second study, we explored how users notice, approach, and engage with public interactive displays in virtual environments. Our aim was to understand if well-known phenomena would persist in VR, such as display blindness [50] and interaction blindness [55] as well as motivation, curiosity, and engagement [46], and the many phases of interaction [44].

Study Design and Task

The study followed a within-subjects design, again the setting (real world / virtual world) being the independent variable. The conditions were counter-balanced across participants. We collected feedback through questionnaires and an interview and we analyzed video recordings of the participants. The study met the ethics regulations of our institution.

Similar to Study 1, we did not want to make it obvious to participants that the public display was under investigation. An additional challenge was to create a situation in which people had the chance to deliberately engage with the display. We addressed this as follows: participants in Study 2 were tasked to interact with an application on a tablet that is in the vicinity of the public display and wait until the experimenter returns. The application was configured to end before the experimenter arrives, thereby creating a *waiting scenario* which gives participants an opportunity to attend to the display.

Apparatus

The public display and the tablet were present in both the real world and VR conditions, and ran an application each.

Public Display Application

We implemented a public display application for use both in the real and virtual worlds (Figure 4). The application was a gesture-controlled game that was controlled via a Kinect One sensor. Users were visualized as skeletons on the display and scored points by catching flying balls. The display showed brief instructions when new users were detected.

Tablet Application

We implemented a tablet application for use in both environments (Figure 4, left). The tablet application was a multiple-choice quiz showing simple questions (e.g., "PEACH is to HCAEP as 46251 is to..." (the correct answer being "15264"). In the real world, the tablet was controlled via touch. In VR, users could play the tablet game by using a controller, with which they could point at and click the buttons on the tablet.

Real-World and VR setups

The real-world condition of Study 2 was conducted in the same location as Study 1. We added the tablet with the quiz application described earlier, and a Kinect One to detect the user's interactions with the public display.

We used the same room model in VR as in Study 1. However, we made several improvements (Figure 1, left & Figure 4). We added textures, improved the lighting and coloring, and added an outside environment populated with trees and bushes, to make the space match its real-world counterpart even better. This way, we also aimed to further ascertain whether the low fidelity of the environment in Study 1 attributed to the overall more active behavior in VR than in the real world.

Unlike in Study 1 where participants had to walk long distances in a virtual space, in Study 2 the available area did not have to be as large. Hence, to allow as much freedom as possible, we used the HTC Vive Pro HMD with a wireless adaptor which allowed participants to move naturally. Another benefit of the wireless upgrade was that users could approach and interact with the public display without obstructions. Still, we used a large meeting room with four trackers (instead of the usual two), so that participants had room to explore the environment in all directions. The borders of the tracking space were visualized inside VR if participants got close to them.

Recruiting

We recruited 16 participants (10 female) through mailing lists. None had participated in Study 1. Their average age was 26.3 (SD = 6.2). Participants were given 30€ as compensation.

Procedure

Results

Our main source of data in Study 2 were the video recordings, which contained the behavior of all participants during both conditions, and the interviews, which were conducted when participants had completed both sessions and the purpose of the study had been revealed to them. Two researchers watched the videos and made notes on participants' behavior and possible interactions with the display. We used inductive content analysis to look for similar themes in the interview responses, and then counted their occurrences.

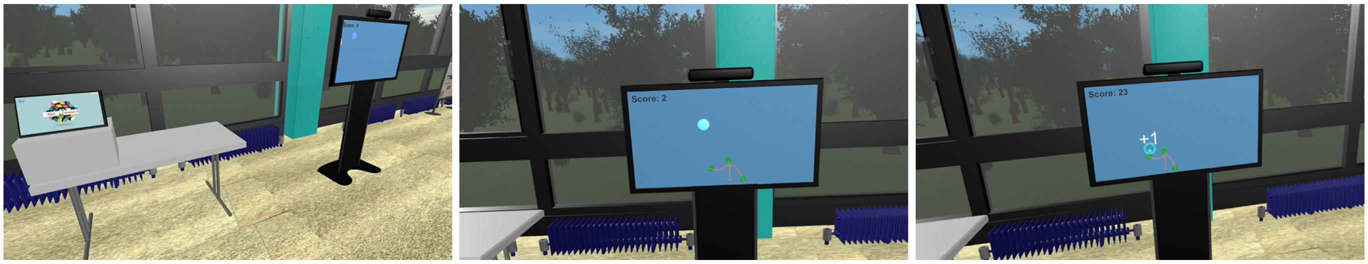

Figure 4. In Study 2, an interactive quiz ran on a tablet (left) and a gesture-controlled game ran on the display (middle), where catch flying balls increased the user's score (right). The experimenter left the participant to interact with the tablet. The tablet's application always ended before the experimenter's return, resulting in a waiting scenario in which the participant had a chance to attend to the display albeit not being instructed to.

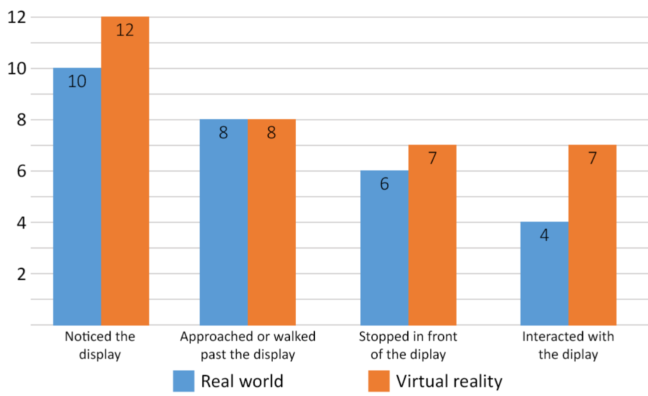

Figure 5. Audience behavior around the virtual and real public displays: how many participants noticed the display, approached or passed the display, stopped in front of the display, and interacted with the display.

Attention towards and Interactions with the Display

We counted the number of participants in both environments with respect to certain actions related to the public display (Figure 5). These were based on the video recordings and the participants' reports during the interview. Overall, the actions are surprisingly similar in both environments. In the real world, 10 participants noticed (looked at) the display, while 12 participants did so in VR. In both environments, eight participants walked past or approached the display. Six participants stopped or stood in front of the display in the real world, seven in VR. A difference was observed in the number of participants interacting with the displays. In VR, seven participants interacted with the display (all of those who stopped in front of it). In the real-world, four participants interacted (four out of the six who stood in front of the display).

It is notable, though, that the two participants in the real world who stood in front of the display but did not interact, seemed to end up in front of the display by accident. In both cases, they were focused on something else, and only noticed the display later when they noticed movement on the screen as well as their skeleton.

Based on the explanations above, 13 out of 16 participants exhibited largely similar behavior in the two environments with respect to their encounters with the displays. The remaining three were those who interacted with the public display in VR but not in the real world, as their behavior in the two environments was noticeably different.

The average interaction time in VR was 109 seconds (SD = 42.1), and the average interaction time in the real world was 130 seconds (SD = 61.3).

Similarities and Differences between the Environments

In the interview, six participants explicitly stated that the environments felt very similar. Others reported various things – big and small – that they noticed or felt were different. Note, that not all of the reported differences necessarily affected the participants' behavior, as some were rather observations or remarks that they came up with when the experiment conductor asked about possible differences between the environments.

Based on both the interview answers and video observations, we identified five unique factors that affected the participants' behavior or experience in VR:

Active Exploration in VR. VR seemed to encourage more exploration than its real-life counterpart. Seven participants self-reported that they were more interested in exploring the VR environment (regardless of whether they felt the environments were similar, or whether they actually explored more in VR). We also observed that participants generally moved more actively in VR and, as previously reported, a greater number of participants interacted with the public display in VR.

Controllers. Participants needed controllers to interact with the tablet game as well as to fill in the questionnaire in VR after the waiting task. Five participants reported that holding the controllers may have affected their behavior, as they were unsure what they could do with them or they feared that they would touch or break something.

Influence of recording. In contrast to users who were more active in exploring the virtual environment than the real environment, five participants reported that the video recording made them less inclined to do anything in VR that they assumed was not part of the study. One participant explained: *"I was conscious and very much aware of recording in VR. I think I did not move too freely"*. Interestingly, though, this was only reported for the VR condition and not the real-world condition. This may be due to the HMD, i.e., participants knew they were being "watched" without being able to see it.

No access to smartphones. In the real world, four participants pulled out their smartphones while waiting and paid little attention to their surroundings. In VR this option did not exist. In fact, one user in VR, likely out of habit, took his smartphone out of the pocket and tried to look at it, but quickly put it away upon realizing that he cannot see it.

Fear of physical harm. Three participants reported that they were wary of moving around in VR, as they feared that something would go wrong or they would bump into something. This seemed to be partly motivated by previous experiences where they had hit something while wearing an HMD.

DISCUSSION

In this section, we discuss the feasibility of conducting field studies in virtual reality by drawing upon the results from our two comparative studies. We discuss the advantages and disadvantages in contrast to traditional lab and field studies, the discovered challenges, and the lessons learned.

User Behavior in Virtual and Real Public Spaces

We observed a surprising level of similarity between virtual and real public spaces in both of our studies. First, in Study 1 we confirmed that the honeypot effect can occur in virtual public spaces as well. Then, in Study 2 we found that people quite similarly notice, approach, and engage with public displays in virtual worlds as they do in real public spaces.

In Study 2, users similarly noticed the public display and understood its interactivity in both environments. This suggests that the early phases of interaction [44] and the related phenomena (e.g., attention [46], display blindness [50], interaction blindness [55]) could be studied in VR. Those who did not notice the display in VR, either did not move much at all or explored the environment in a direction away from the display. In contrast, those who did not notice the display in the real world, usually pulled out their smartphone and/or sat on a bench. Nonetheless, the result from a display effectiveness standpoint was similar. These attention-related conclusions are further supported by our findings in Study 1, where the honeypot effect resulted not only in increased gaze duration, but also an increase in how many participants remembered the content on the display.

Interaction in Study 2 also seemed similar. Interaction times were in the same order of magnitude (close to two minutes) in both environments. Based on the video recordings, users were visibly excited about the public display, actively moving from left to right and reaching with their arms to catch as many balls as possible. This clearly links to playful behavior that has been observed many times in front of real-world displays, especially gesture-controlled displays [2, 17, 47, 48, 76]. An exception was a user who interacted in a very subtle manner using only one hand, but even in this case the participant's behavior was the same with both displays.

Despite many similarities, the behaviors are not entirely matching, which provides some limitations and uncertainties as to what conclusions we can draw from virtual field studies. Most importantly, *people are more interested in virtual environments than in real public spaces*. We observed several implications as a result of this, the primary one of which is that people may be more likely to interact with a virtual public display than a real public display (7 and 4 users, respectively). Therefore, it is unclear if VR can be used to study motivation or encouragement [46]. This finding also suggests that the lower fidelity of the virtual environment in Study 1 was not a significant factor in drawing attention to the public display, since we observed similarly more active behavior in VR in Study 2.

Virtual Field Studies - A New Research Approach

At the core of our investigation is **A)** whether virtual field studies can produce valid results (i.e., results that would be similar to the results from a real lab or field study), and **B)** whether virtual field studies can be conducted with low enough effort that its use in some studies can be argued for.

Our early investigation suggests that it is indeed possible to produce valid results with virtual field studies. We believe that *virtual field studies can reach high ecological validity* for certain types of research questions, such as those focused on audience behavior or display effectiveness [9]. Still, we uncovered some aspects and challenges unique to VR. It is therefore likely that virtual field studies will not reach the ecological validity of traditional field studies, at least not for all types of studies.

As for the effort, we argue that digital deployments are often cheaper and faster to build than real deployments. Building the virtual environment itself can take some time. However, modern tools and assets – many of which are freely available – have made this process significantly easier. Most importantly, however, virtual field studies are not subject to the many difficulties that accompany traditional field studies, such as permissions and maintenance [5, 43, 54, 72, 75]. Still, for individual studies with a single display configuration, building a virtual environment may not always be worth the effort.

Based on our results and the discussion above, we introduce virtual field studies as a new research paradigm (Figure 6). In terms of required effort and ecological validity, we argue that virtual field studies situate between traditional lab and field studies, being closer to field studies in ecological validity and closer to lab studies in effort, but with some overlap with both.

Lessons Learned

Based on our study results as well as on our experiences designing and conducting two virtual field studies, we distill lessons learned to help researchers and practitioners consider the use of virtual field studies, and to avoid potential pitfalls.

Lesson 1: When to Consider Virtual Field Studies

We believe that virtual field studies can be especially valuable when one or more of the following apply: **A)** The study contains several conditions. In particular, VR makes it easy to experiment with different physical display configurations, e.g., layouts, shapes, sizes and locations. **B)** The same virtual configuration (or parts of it) will be used for several studies. Real-world setups are often difficult and expensive to maintain [43, 54, 72], whereas VR setups do not need similar attention. **C)** The targeted real-world space is not available. This may be due to the uncooperative nature of different stakeholders such as location managers [24, 32] or difficulties with obtaining permissions for collecting data [5, 75], or perhaps the space does not yet exist (e.g., it is under construction).

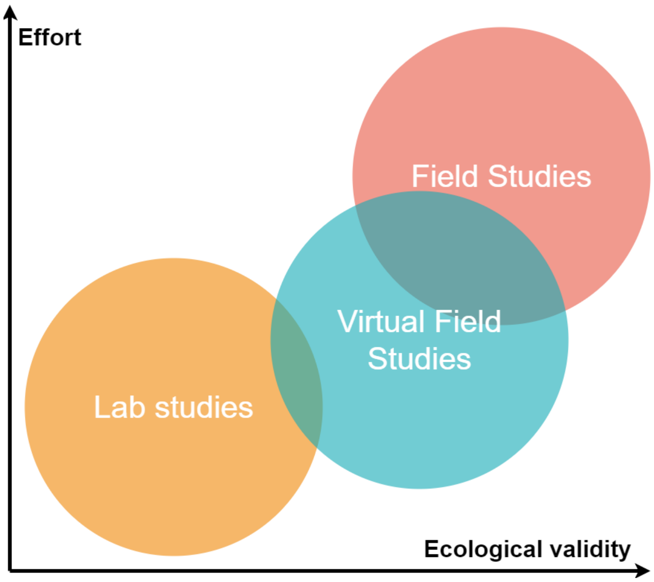

Figure 6. We introduce virtual field studies as a new research paradigm. We estimate that VR field studies situate between lab studies and real-world field studies, being closer to field studies in ecological validity, and closer to lab studies with regards to the required effort.

Still, there are types of studies and situations where real-world studies (either in the field or in the lab) are likely more suitable. Such situations include, but are not limited to: **A)** The display setup is already prepared. **B)** There are environmental aspects that cannot be replicated in VR. **C)** The study focuses on long-term or continuous use. Getting users into a virtual space on a recurring basis is a challenge in VR studies.

There are also other studies and situations where potential pitfalls exist with regards to virtual field studies, or where we lack knowledge to make sufficient conclusions on the feasibility of virtual field studies. However, many of these aspects can be overcome, and more knowledge can be gained through further research. Therefore, we discuss these aspects in the future work section instead.

Lesson 2: Provide Context for VR Users

In real public spaces, people have inherent needs and motivations for being at the location, such as passing through, waiting for or meeting with someone, studying, or having lunch. People rarely deliberately seek out public displays, but rather, they stumble upon them and use them on an opportunity basis [46].

In contrast, putting on an HMD and immersing oneself in a virtual world is a significantly different setting. As such, for conducting field studies in VR, users must be given a context or task that **A)** simulates a natural scenario, **B)** puts users (but does not force them!) in an ideal situation where they have the chance to engage in the measured activity, and **C)** is not in conflict with the measured activity.

To explain this reasoning in more detail, we review the designs of our two studies against these criteria:

In Study 1, the participants' task had them move through the space, simulating a "passing through" situation **(A)**. This put them in an ideal situation where they passed by the public display and could observe it if they wanted to **(B)**. Looking at the display was not in conflict with the task, as looking had little to no effect on the task performance **(C)**. As a counter example, if we had used the same task but instead measured, say, *interaction* with the display, then there would have been a conflict, because the ongoing task put pressure on participants to keep moving and complete the task in a speedy manner, influencing their willingness to stay and explore the display.

Applying the same criteria for Study 2, the participants' task put them in a "waiting for someone" situation **(A)**. This gave them an opportunity to freely move around the space and possibly stumble upon the public display **(B)**. Exploring the space and possibly the display were not in conflict with the task, since participants knew they were waiting and that they would know when the waiting is over **(C)**.

Lesson 3: Be Wary of Gaze and Movement Data

We learned from our studies that people are more interested in virtual environments than in real public spaces, making them move more and look at things for longer. Hence, while we saw in both studies that overall attention and behavior matches quite well, we also saw that detailed analysis using gaze or other movement data may be unreliable. This was particularly evident in Study 1, where VR users' attention towards the public display was several times longer than in the real world.

We do not outright dismiss the value of such data. After all, we used gaze data to confirm the existence of the honeypot effect in Study 1; however, this was mainly used in both environments separately, comparing two conditions (audience and no audience) within both of them – not as a direct comparison between a virtual and a real public space. We further complemented this analysis by also asking participants to report the contents of the display, as additional proof of their focus.

Our takeaway is that such quantitative data should be considered carefully, and in-depth quantitative comparisons between virtual and public spaces should be avoided until further research provides a better understanding on the detailed behavioral differences between virtual and real public spaces.

Future Work

In this paper, we have taken the first steps towards understanding the capabilities and challenges of using VR as a research approach. Still, we acknowledge that more studies are required before making generalized conclusions about people's behavior in virtual environments, and before reaching an adequate understanding of using VR for user studies.

To this end, we recognize three important areas that should be studied in the future: **1)** social encounters (e.g., the influence of crowds and people), **2)** direct interaction with virtual displays and different interaction modalities, **3)** non-visual aspects of virtual environments. We briefly discuss these areas as follows:

First, a defining characteristic of public spaces is that they are actively populated by people. Past research has shown that nearby people have considerable influence on whether a person is attentive to the public display and whether they approach or interact with it [13, 16, 39, 63]. While we did touch on this subject by studying the honeypot effect, many significant questions remain. In particular, it would be valuable to investigate whether the presence of virtual characters can induce similar feelings as real people, such as unwillingness to interact with a public display when someone is watching [61]. Similarly, it would be worthwhile to observe how groups of users engage with virtual public displays and how they interact with each other. Such situations could be complemented with an audio channel to allow for deeper social interactions.

Second, our investigation was not fully focused on direct interaction [44] with virtual displays. In the first study, participants had no direct interactions with the public display and in the second study, we focused on measuring interaction times and overall behavior through video observations. Hence, there is need for an in-depth exploration of this phase of interaction in VR, in particular with regards to various interaction modalities. An obvious investigation would be to look into traditional touchscreens in VR, which could be implemented by using trackers to map virtual displays to physical displays to enable natural haptic feedback. Such solutions would also avoid potential pitfalls with handheld controllers in VR; the results of our second study suggested the controllers somewhat affected the behavior of some participants.

Third, our work was focused on visual feedback when building the virtual space, displays, and other relevant characteristics. Real urban environments can be noisy, which can often direct people's attention and distract them. We therefore predict an increase in ecological validity if this aspect is considered in future studies, for example by adding natural audio sources and ambience in the virtual environment.

CONCLUSION

Public display research is often conducted in the field in realistic settings. However, field studies are often difficult, expensive, and time-consuming. Moreover, field studies provide limited insight relative to the required workload, as they explore a specific setting and a small number of display configurations.

In this paper, we have introduced the concept of virtual field studies: conducting field studies on public displays in virtual reality (VR). Virtual field studies offer many potential benefits over traditional field studies. Most importantly, VR allows for practically infinite control over the setting and the display deployment, allowing for quick testing of different display configurations, physical layouts, and situations.

To evaluate the feasibility of conducting field studies in VR, our first step was to investigate how well results gained from virtual field studies would compare to results from real-world studies. To this end, we conducted two user studies, where we compared user behavior in a real public space to user behavior in a comparable virtual space. In the first study, we confirmed that one of the most well-known phenomena in public display research, the honeypot effect, can occur in VR. In the second study, we studied how users notice, approach, and interact with public displays in VR. We observed surprising similarity with real-world behaviors: users similarly took notice and approached the public display. Their way of interacting with the displays were also similar. A notable difference was observed in that VR users were more motivated to explore, also resulting in more users interacting with the public display in VR.

Through this research, we have taken the first steps towards understanding the potential, strengths, and weaknesses of using VR as a tool for conducting field studies. Our work suggests that virtual field studies can be a powerful research approach for evaluating public displays and potentially other technologies, but certain considerations must be accounted for, as discussed in this paper.

ACKNOWLEDGMENTS

This work was funded by the Ulla Tuominen Foundation and the Foundation's Post Doc Pool as well as by the Royal Society of Edinburgh (Award number 65040), the Deutsche Forschungsgemeinschaft (DFG) (Grant numbers AL 1899/2-1 and 1899/4-1), and the European Union's Horizon 2020 Programme under ERCEA grant no. 683008 AMPLIFY.

REFERENCES

[1] Christopher Ackad, Andrew Clayphan, Martin Tomitsch, and Judy Kay. 2015. An In-the-wild Study of Learning Mid-air Gestures to Browse Hierarchical Information at a Large Interactive Public Display. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 1227–1238. DOI: http://dx.doi.org/10.1145/2750858.2807532

[2] Christopher Ackad, Martin Tomitsch, and Judy Kay. 2016. Skeletons and Silhouettes: Comparing User Representations at a Gesture-based Large Display. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2343–2347. DOI: http://dx.doi.org/10.1145/2858036.2858427

[3] Christopher Ackad, Rainer Wasinger, Richard Gluga, Judy Kay, and Martin Tomitsch. 2013. Measuring Interactivity at an Interactive Public Information Display. In *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration (OzCHI '13)*. ACM, New York, NY, USA, 329–332. DOI: http://dx.doi.org/10.1145/2541016.2541091

[4] Philipp Agethen, Viswa Subramanian Sekar, Felix Gaisbauer, Thies Pfeiffer, Michael Otto, and Enrico Rukzio. 2018. Behavior Analysis of Human Locomotion in the Real World and Virtual Reality for the Manufacturing Industry. *ACM Trans. Appl. Percept.* 15, 3, Article 20 (July 2018), 19 pages. DOI: http://dx.doi.org/10.1145/3230648

[5] Jan Philipp Albrecht. 2016. How the GDPR Will Change the World. *European Data Protection Law Review* 2, 3 (2016). DOI:http://dx.doi.org/10.21552/EDPL/2016/3/4

[6] Florian Alt, Andreas Bulling, Gino Gravanis, and Daniel Buschek. 2015. GravitySpot: Guiding Users in Front of Public Displays Using On-Screen Visual Cues. In *Proceedings of the 28th Annual ACM Symposium on*

User Interface Software & Technology (UIST '15). ACM, New York, NY, USA, 47–56. DOI: http://dx.doi.org/10.1145/2807442.2807490

[7] Florian Alt, Andreas Bulling, Lukas Mecke, and Daniel Buschek. 2016. Attention, Please! Comparing Features for Measuring Audience Attention Towards Pervasive Displays. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. Association for Computing Machinery, New York, NY, USA, 823–828. DOI: http://dx.doi.org/10.1145/2901790.2901897

[8] Florian Alt, Thomas Kubitza, Dominik Bial, Firas Zaidan, Markus Ortel, Björn Zurmaar, Tim Lewen, Alireza Sahami Shirazi, and Albrecht Schmidt. 2011. Digifieds: Insights into Deploying Digital Public Notice Areas in the Wild. In *Proceedings of the 10th International Conference on Mobile and Ubiquitous Multimedia (MUM '11)*. ACM, New York, NY, USA, 165–174. DOI: http://dx.doi.org/10.1145/2107596.2107618

[9] Florian Alt, Stefan Schneegaß, Albrecht Schmidt, Jörg Müller, and Nemanja Memarovic. 2012. How to Evaluate Public Displays. In *Proceedings of the 2012 International Symposium on Pervasive Displays (PerDis '12)*. ACM, New York, NY, USA, Article 17, 6 pages. DOI:http://dx.doi.org/10.1145/2307798.2307815

[10] Florian Alt, Alireza Sahami Shirazi, Thomas Kubitza, and Albrecht Schmidt. 2013. Interaction Techniques for Creating and Exchanging Content with Public Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1709–1718. DOI: http://dx.doi.org/10.1145/2470654.2466226

[11] Florian Alt and Julia Vehns. 2016. Opportunistic Deployments: Challenges and Opportunities of Conducting Public Display Research at an Airport. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays (PerDis '16)*. Association for Computing Machinery, New York, NY, USA, 106–117. DOI:http://dx.doi.org/10.1145/2914920.2915020

[12] John J Barton and Vikram Vijayaraghavan. 2002. Ubiwise, a ubiquitous wireless infrastructure simulation environment. *HP Labs* (2002).

[13] Ben Bedwell and Theresa Caruana. 2012. Encouraging Spectacle to Create Self-sustaining Interactions at Public Displays. In *Proceedings of the 2012 International Symposium on Pervasive Displays (PerDis '12)*. ACM, New York, NY, USA, Article 15, 6 pages. DOI: http://dx.doi.org/10.1145/2307798.2307813

[14] Gilbert Beyer, Florian Alt, Jörg Müller, Albrecht Schmidt, Karsten Isakovic, Stefan Klose, Manuel Schiewe, and Ivo Haulsen. 2011. Audience Behavior Around Large Interactive Cylindrical Screens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 1021–1030. DOI: http://dx.doi.org/10.1145/1978942.1979095

[15] Gilbert Beyer, Vincent Binder, Nina Jäger, and Andreas Butz. 2014. The Puppeteer Display: Attracting and Actively Shaping the Audience with an Interactive Public Banner Display. In *Proceedings of the 2014 Conference on Designing Interactive Systems (DIS '14)*. ACM, New York, NY, USA, 935–944. DOI: http://dx.doi.org/10.1145/2598510.2598575

[16] Harry Brignull and Yvonne Rogers. 2003. Enticing People to Interact with Large Public Displays in Public Spaces. In *Human-Computer Interaction INTERACT '03: IFIP TC13 International Conference on Human-Computer Interaction, 1st-5th September 2003, Zurich, Switzerland*.

[17] Peter Dalsgaard and Kim Halskov. 2010. Designing Urban Media FaçAdes: Cases and Challenges. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 2277–2286. DOI: http://dx.doi.org/10.1145/1753326.1753670

[18] Peter Dalsgaard, Kim Halskov, and Ole Sejer Iversen. 2016. Participation Gestalt: Analysing Participatory Qualities of Interaction in Public Space. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4435–4446. DOI: http://dx.doi.org/10.1145/2858036.2858147

[19] Peter Dalsgaard and Lone Koefoed Hansen. 2008. Performing Perception—Staging Aesthetics of Interaction. *ACM Trans. Comput.-Hum. Interact.* 15, 3, Article 13 (Dec. 2008), 33 pages. DOI: http://dx.doi.org/10.1145/1453152.1453156

[20] Nicholas S. Dalton, Emily Collins, and Paul Marshall. 2015. Display Blindness?: Looking Again at the Visibility of Situated Displays Using Eye-tracking. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3889–3898. DOI: http://dx.doi.org/10.1145/2702123.2702150

[21] Nigel Davies, Sarah Clinch, and Florian Alt. 2014. *Pervasive Displays: Understanding the Future of Digital Signage* (1st ed.). Morgan & Claypool Publishers.

[22] Shuchisnigdha Deb, Daniel W. Carruth, Richard Sween, Lesley Strawderman, and Teena M. Garrison. 2017. Efficacy of virtual reality in pedestrian safety research. *Applied Ergonomics* 65 (2017), 449 – 460. DOI: http://dx.doi.org/10.1016/j.apergo.2017.03.007

[23] Guiying Du, Auriol Degbelo, and Christian Kray. 2017. Public Displays for Public Participation in Urban Settings: A Survey. In *Proceedings of the 6th ACM International Symposium on Pervasive Displays (PerDis '17)*. ACM, New York, NY, USA, Article 17, 9 pages. DOI:http://dx.doi.org/10.1145/3078810.3078825

[24] Ivan Elhart, Marc Langheinrich, Nemanja Memarovic, and Elisa Rubegni. 2016. A Good Balance of Costs and Benefits: Convincing a University Administration to Support the Installation of an Interactive Multi-application Display System on Campus. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays (PerDis '16)*. ACM, New York, NY, USA, 197–203. DOI: http://dx.doi.org/10.1145/2914920.2915029

[25] Maia Garau, Mel Slater, Vinoba Vinayagamoorthy, Andrea Brogni, Anthony Steed, and M. Angela Sasse. 2003. The Impact of Avatar Realism and Eye Gaze Control on Perceived Quality of Communication in a Shared Immersive Virtual Environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 529–536. DOI: http://dx.doi.org/10.1145/642611.642703

[26] Vito Gentile, Mohamed Khamis, Salvatore Sorce, and Florian Alt. 2017. They Are Looking at Me!: Understanding How Audience Presence Impacts on Public Display Users. In *Proceedings of the 6th ACM International Symposium on Pervasive Displays (PerDis '17)*. ACM, New York, NY, USA, Article 11, 7 pages. DOI:http://dx.doi.org/10.1145/3078810.3078822

[27] Dominic Gorecky, Mohamed Khamis, and Katharina Mura. 2017. Introduction and Establishment of Virtual Training in the Factory of the Future. *International Journal on Computer Integrated Manufacturing* 30, 1 (Jan. 2017), 182–190. DOI: http://dx.doi.org/10.1080/0951192X.2015.1067918

[28] John Hardy, Enrico Rukzio, and Nigel Davies. 2011. Real World Responses to Interactive Gesture Based Public Displays. In *Proceedings of the 10th International Conference on Mobile and Ubiquitous Multimedia (MUM '11)*. ACM, New York, NY, USA, 33–39. DOI:http://dx.doi.org/10.1145/2107596.2107600

[29] Eric Harris, Geraldine Fitzpatrick, Yvonne Rogers, Sara Price, Ted Phelps, and Cliff Randell. 2004. From Snark to Park: Lessons Learnt Moving Pervasive Experiences from Indoors to Outdoors. In *Proceedings of the Fifth Conference on Australasian User Interface - Volume 28 (AUIC '04)*. Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 39–48. http://dl.acm.org/citation.cfm?id=976310.976316

[30] Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage publications Sage CA: Los Angeles, CA, 904–908. DOI: http://dx.doi.org/10.1177/154193120605000909

[31] Marc Hassenzahl, Michael Burmester, and Franz Koller. 2003. *AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität*. Vieweg+Teubner Verlag, Wiesbaden, 187–196. DOI:http://dx.doi.org/10.1007/978-3-322-80058-9_19

[32] Simo Hosio, Jorge Goncalves, Hannu Kukka, Alan Chamberlain, and Alessio Malizia. 2014. What's in It for Me: Exploring the Real-World Value Proposition of Pervasive Displays. In *Proceedings of The International Symposium on Pervasive Displays (PerDis '14)*. ACM, New York, NY, USA, Article 174, 6 pages. DOI: http://dx.doi.org/10.1145/2611009.2611012

[33] Simo Hosio, Vassilis Kostakos, Hannu Kukka, Marko Jurmu, Jukka Riekki, and Timo Ojala. 2012. From School Food to Skate Parks in a Few Clicks: Using Public Displays to Bootstrap Civic Engagement of the Young. In *Pervasive Computing*, Judy Kay, Paul Lukowicz, Hideyuki Tokuda, Patrick Olivier, and Antonio Krüger (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 425–442.

[34] Arief Ernst Hühn, Vassilis-Javed Khan, Andrés Lucero, and Paul Ketelaar. 2012. On the Use of Virtual Environments for the Evaluation of Location-based Applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 2569–2578. DOI: http://dx.doi.org/10.1145/2207676.2208646

[35] Wendy Ju and David Sirkin. 2010. Animate Objects: How Physical Motion Encourages Public Interaction. In *Persuasive Technology*, Thomas Ploug, Per Hasle, and Harri Oinas-Kukkonen (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 40–51.

[36] Mohamed Khamis, Florian Alt, and Andreas Bulling. 2015. A Field Study on Spontaneous Gaze-based Interaction with a Public Display Using Pursuits. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC'15 Adjunct)*. ACM, New York, NY, USA, 863–872. DOI: http://dx.doi.org/10.1145/2800835.2804335

[37] Mohamed Khamis, Christian Becker, Andreas Bulling, and Florian Alt. 2018. Which One is Me?: Identifying Oneself on Public Displays. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 287, 12 pages. DOI: http://dx.doi.org/10.1145/3173574.3173861

[38] Hannu Kukka, Heidi Oja, Vassilis Kostakos, Jorge Gonçalves, and Timo Ojala. 2013. What Makes You Click: Exploring Visual Signals to Entice Interaction on Public Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1699–1708. DOI:http://dx.doi.org/10.1145/2470654.2466225

[39] Andrés Lucero, Jussi Holopainen, and Tero Jokela. 2012. MobiComics: Collaborative Use of Mobile Phones and Large Displays for Public Expression. In *Proceedings of the 14th International Conference on Human-computer*

Interaction with Mobile Devices and Services (MobileHCI '12). ACM, New York, NY, USA, 383–392. DOI:http://dx.doi.org/10.1145/2371574.2371634

[40] Ville Mäkelä, Tomi Heimonen, Matti Luhtala, and Markku Turunen. 2014. Information Wall: Evaluation of a Gesture-controlled Public Display. In *Proceedings of the 13th International Conference on Mobile and Ubiquitous Multimedia (MUM '14)*. ACM, New York, NY, USA, 228–231. DOI: http://dx.doi.org/10.1145/2677972.2677998

[41] Paul Marshall, Richard Morris, Yvonne Rogers, Stefan Kreitmayer, and Matt Davies. 2011. Rethinking 'Multi-user': An In-the-wild Study of How Groups Approach a Walk-up-and-use Tabletop Interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 3033–3042. DOI: http://dx.doi.org/10.1145/1978942.1979392

[42] Nemanja Memarovic, Sarah Clinch, and Florian Alt. 2015. Understanding Display Blindness in Future Display Deployments. In *Proceedings of the 4th International Symposium on Pervasive Displays (PerDis '15)*. ACM, New York, NY, USA, 7–14. DOI: http://dx.doi.org/10.1145/2757710.2757719

[43] Nemanja Memarovic, Marc Langheinrich, Keith Cheverst, Nick Taylor, and Florian Alt. 2013. P-LAYERS – A Layered Framework Addressing the Multifaceted Issues Facing Community-Supporting Public Display Deployments. *ACM Trans. Comput.-Hum. Interact.* 20, 3, Article 17 (July 2013), 34 pages. DOI:http://dx.doi.org/10.1145/2491500.2491505

[44] Daniel Michelis and Jörg Müller. 2011. The Audience Funnel: Observations of Gesture Based Interaction With Multiple Large Displays in a City Center. *International Journal of Human–Computer Interaction* 27, 6 (2011), 562–579. DOI: http://dx.doi.org/10.1080/10447318.2011.555299

[45] Mehdi Moussaïd, Mubbasir Kapadia, Tyler Thrash, Robert W Sumner, Markus Gross, Dirk Helbing, and Christoph Hölscher. 2016. Crowd behaviour during high-stress evacuations in an immersive virtual environment. *Journal of The Royal Society Interface* 13, 122 (2016), 20160414.

[46] Jörg Müller, Florian Alt, Daniel Michelis, and Albrecht Schmidt. 2010. Requirements and Design Space for Interactive Public Displays. In *Proceedings of the 18th ACM International Conference on Multimedia (MM '10)*. ACM, New York, NY, USA, 1285–1294. DOI: http://dx.doi.org/10.1145/1873951.1874203

[47] Jörg Müller, Gilles Bailly, Thor Bossuyt, and Niklas Hillgren. 2014a. MirrorTouch: Combining Touch and Mid-air Gestures for Public Displays. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 319–328. DOI: http://dx.doi.org/10.1145/2628363.2628379

[48] Jörg Müller, Dieter Eberle, and Konrad Tollmar. 2014b. Communiplay: A Field Study of a Public Display Mediaspace. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1415–1424. DOI:http://dx.doi.org/10.1145/2556288.2557001

[49] Jörg Müller, Robert Walter, Gilles Bailly, Michael Nischt, and Florian Alt. 2012. Looking Glass: A Field Study on Noticing Interactivity of a Shop Window. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 297–306. DOI: http://dx.doi.org/10.1145/2207676.2207718

[50] Jörg Müller, Dennis Wilmsmann, Juliane Exeler, Markus Buzeck, Albrecht Schmidt, Tim Jay, and Antonio Krüger. 2009. Display Blindness: The Effect of Expectations on Attention towards Digital Signage. In *Pervasive Computing*, Hideyuki Tokuda, Michael Beigl, Adrian Friday, A. J. Bernheim Brush, and Yoshito Tobe (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 1–8.

[51] Maximilian Müller, Aris Alissandrakis, and Nuno Otero. 2016. There is More to Come: Anticipating Content on Interactive Public Displays Through Timer Animations. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays (PerDis '16)*. ACM, New York, NY, USA, 247–248. DOI: http://dx.doi.org/10.1145/2914920.2940341

[52] Ville Mäkelä, Tomi Heimonen, and Markku Turunen. 2018a. Semi-Automated, Large-Scale Evaluation of Public Displays. *International Journal of Human–Computer Interaction* 34, 6 (2018), 491–505. DOI:http://dx.doi.org/10.1080/10447318.2017.1367905

[53] Ville Mäkelä, Mohamed Khamis, Lukas Mecke, Jobin James, Markku Turunen, and Florian Alt. 2018b. Pocket Transfers: Interaction Techniques for Transferring Content from Situated Displays to Mobile Devices. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 135, 13 pages. DOI: http://dx.doi.org/10.1145/3173574.3173709

[54] Ville Mäkelä, Sumita Sharma, Jaakko Hakulinen, Tomi Heimonen, and Markku Turunen. 2017. Challenges in Public Display Deployments: A Taxonomy of External Factors. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 3426–3475. DOI: http://dx.doi.org/10.1145/3025453.3025798

[55] T. Ojala, V. Kostakos, H. Kukka, T. Heikkinen, T. Linden, M. Jurmu, S. Hosio, F. Kruger, and D. Zanni. 2012. Multipurpose Interactive Public Displays in the Wild: Three Years Later. *Computer* 45, 5 (May 2012), 42–49. DOI:http://dx.doi.org/10.1109/MC.2012.115

[56] T. Ojala, H. Kukka, T. Lindén, T. Heikkinen, M. Jurmu, S. Hosio, and F. Kruger. 2010. UBI-Hotspot 1.0: Large-Scale Long-Term Deployment of Interactive Public Displays in a City Center. In *2010 Fifth International Conference on Internet and Web Applications and Services*. 285–294. DOI: http://dx.doi.org/10.1109/ICIW.2010.49

[57] Néstor Ordaz, David Romero, Dominic Gorecky, and Héctor R. Siller. 2015. Serious Games and Virtual Simulator for Automotive Manufacturing Education & Training. *Procedia Computer Science* 75 (2015), 267 – 274. DOI: http://dx.doi.org/10.1016/j.procs.2015.12.247

[58] Callum Parker and Martin Tomitsch. 2017. Bridging the Interaction Gulf: Understanding the Factors That Drive Public Interactive Display Usage. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction (OZCHI '17)*. ACM, New York, NY, USA, 482–486. DOI: http://dx.doi.org/10.1145/3152771.3156162

[59] Gonzalo Parra, Robin De Croon, Joris Klerkx, and Erik Duval. 2014. Quantifying the Interaction Stages of a Public Display Campaign in the Wild. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI '14)*. ACM, New York, NY, USA, 757–760. DOI: http://dx.doi.org/10.1145/2639189.2639216

[60] Peter Peltonen, Esko Kurvinen, Antti Salovaara, Giulio Jacucci, Tommi Ilmonen, John Evans, Antti Oulasvirta, and Petri Saarikko. 2008. It's Mine, Don'T Touch!: Interactions at a Large Multi-touch Display in a City Centre. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 1285–1294. DOI: http://dx.doi.org/10.1145/1357054.1357255

[61] Mark Perry, Steve Beckett, Kenton O'Hara, and Sriram Subramanian. 2010. WaveWindow: Public, Performative Gestural Interaction. In *ACM International Conference on Interactive Tabletops and Surfaces (ITS '10)*. ACM, New York, NY, USA, 109–112. DOI: http://dx.doi.org/10.1145/1936652.1936672

[62] Sharif Razzaque, David Swapp, Mel Slater, Mary C. Whitton, and Anthony Steed. 2002. Redirected Walking in Place. In *Proceedings of the Workshop on Virtual Environments 2002 (EGVE '02)*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 123–130. http://dl.acm.org/citation.cfm?id=509709.509729

[63] Stuart Reeves. 2011. *Designing interfaces in public settings: Understanding the role of the spectator in Human-Computer Interaction*. Springer Science & Business Media.

[64] W. Reitberger, A. Meschtscherjakov, T. Mirlacher, T. Scherndl, H. Huber, and M. Tscheligi. 2009. A Persuasive Interactive Mannequin for Shop Windows. In *Proceedings of the 4th International Conference on Persuasive Technology (Persuasive '09)*. ACM, New York, NY, USA, Article 4, 8 pages. DOI: http://dx.doi.org/10.1145/1541948.1541954

[65] A. Riener and A. Sippl. 2014. Head-Pose-Based Attention Recognition on Large Public Displays. *IEEE Computer Graphics and Applications* 34, 1 (Jan 2014), 32–41. DOI: http://dx.doi.org/10.1109/MCG.2014.9

[66] Enrico Rukzio, Michael Müller, and Robert Hardy. 2009. Design, Implementation and Evaluation of a Novel Public Display for Pedestrian Navigation: The Rotating Compass. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 113–122. DOI: http://dx.doi.org/10.1145/1518701.1518722

[67] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2017. Construction of a Benchmark for the User Experience Questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence* 4, 4 (06/2017 2017), 40–44. DOI: http://dx.doi.org/10.9781/ijimai.2017.445

[68] Helmut Schrom-Feiertag, Christoph Schinko, Volker Settgast, and Stefan Seer. 2014. Evaluation of Guidance Systems in Public Infrastructures Using Eye Tracking in an Immersive Virtual Environment. *CEUR Workshop Proceedings* 1241 (01 2014), 62–66. http://ceur-ws.org/Vol-1241/paper13.pdf

[69] Helmut Schrom-Feiertag, Volker Settgast, and Stefan Seer. 2017. Evaluation of indoor guidance systems using eye tracking in an immersive virtual environment. *Spatial Cognition & Computation* 17, 1-2 (2017), 163–183.

[70] Valentin Schwind, Pascal Knierim, Nico Haas, and Niels Henze. 2019. Using Presence Questionnaires in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 360, 12 pages. DOI: http://dx.doi.org/10.1145/3290605.3300590

[71] Srinivas Sridharan, Reynold Bailey, Ann McNamara, and Cindy Grimm. 2012. Subtle Gaze Manipulation for Improved Mammography Training. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12)*. ACM, New York, NY, USA, 75–82. DOI: http://dx.doi.org/10.1145/2168556.2168568

[72] O. Storz, A. Friday, N. Davies, J. Finney, C. Sas, and J. Sheridan. 2006. Public Ubiquitous Computing Systems: Lessons from the e-Campus Display Deployments. *IEEE Pervasive Computing* 5, 3 (July 2006), 40–47. DOI: http://dx.doi.org/10.1109/MPRV.2006.56

[73] Maurice Ten Koppel, Gilles Bailly, Jörg Müller, and Robert Walter. 2012a. Chained Displays: Configurations of Public Displays Can Be Used to Influence Actor-, Audience-, and Passer-by Behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 317–326. DOI: http://dx.doi.org/10.1145/2207676.2207720

[74] Maurice Ten Koppel, Gilles Bailly, Jörg Müller, and Robert Walter. 2012b. Chained Displays: Configurations of Public Displays Can Be Used to Influence Actor-, Audience-, and Passer-by Behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 317–326. DOI: http://dx.doi.org/10.1145/2207676.2207720

[75] Eva Thelisson, Kshitij Sharma, Hanan Salam, and Virginia Dignum. 2018. The General Data Protection Regulation: An Opportunity for the HCI Community?. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, Article Paper W36, 8 pages. DOI: http://dx.doi.org/10.1145/3170427.3170632

[76] Martin Tomitsch, Christopher Ackad, Oliver Dawson, Luke Hespanhol, and Judy Kay. 2014. Who Cares About the Content? An Analysis of Playful Behaviour at a Public Display. In *Proceedings of The International Symposium on Pervasive Displays (PerDis '14)*. ACM, New York, NY, USA, Article 160, 6 pages. DOI: http://dx.doi.org/10.1145/2611009.2611016

[77] Sam Tregillus and Eelke Folmer. 2016. VR-STEP: Walking-in-Place Using Inertial Sensing for Hands Free Navigation in Mobile VR Environments. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1250–1255. DOI: http://dx.doi.org/10.1145/2858036.2858084

[78] Paco Underhill. 2009. *Why we buy: The science of shopping–updated and revised for the Internet, the global consumer, and beyond*. Simon and Schuster.

[79] Nina Valkanova, Robert Walter, Andrew Vande Moere, and Jörg Müller. 2014. MyPosition: Sparking Civic Discourse by a Public Interactive Poll Visualization. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 1323–1332. DOI:http://dx.doi.org/10.1145/2531602.2531639

[80] Robert Walter, Gilles Bailly, and Jörg Müller. 2013. StrikeAPose: Revealing Mid-air Gestures on Public Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 841–850. DOI: http://dx.doi.org/10.1145/2470654.2470774

[81] Robert Walter, Andreas Bulling, David Lindlbauer, Martin Schuessler, and Jörg Müller. 2015. Analyzing Visual Attention During Whole Body Interaction with Public Displays. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 1263–1267. DOI: http://dx.doi.org/10.1145/2750858.2804255

[82] Julie R. Williamson and Daniel Sundén. 2015. Enter the Circle: Blending Spherical Displays and Playful Embedded Interaction in Public Spaces. In *Proceedings of the 4th International Symposium on Pervasive Displays (PerDis '15)*. ACM, New York, NY, USA, 195–200. DOI: http://dx.doi.org/10.1145/2757710.2757731

[83] Niels Wouters, John Downs, Mitchell Harrop, Travis Cox, Eduardo Oliveira, Sarah Webber, Frank Vetere, and Andrew Vande Moere. 2016. Uncovering the Honeypot Effect: How Audiences Engage with Public Interactive Systems. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. ACM, New York, NY, USA, 5–16. DOI: http://dx.doi.org/10.1145/2901790.2901796