Who cares about the Content? An Analysis of Playful Behaviour at a Public Display

Martin Tomitsch¹, Christopher Ackad², Oliver Dawson¹,², Luke Hespanhol¹, Judy Kay²
¹ Faculty of Architecture, Design and Planning – Design Lab
² School of Information Technologies – Computer Human Adapted Interaction (CHAI)
The University of Sydney, Sydney NSW 2006, Australia
{firstname.lastname}@sydney.edu.au

ABSTRACT
In this paper, we report on a field deployment study of a public interactive display, in which we observed a surprising number of interactions that seemed to be more concerned about playing ‘with’ the display rather than exploring its content. The display featured information about events at a nearby theatre and activities at the university, and supported four basic gestures for navigating through the content. To indicate its interactive capabilities, the display represented passers-by as a mirror image in the form of a skeleton. Our analysis of depth video recordings suggests that this representation may have triggered some of the playful behaviour we observed in the deployment study. To better understand how and when people engaged in playful behaviours, we conducted an in-depth analysis of the 40 recordings of longest duration. These had a total of 102 people recorded over an 8-day period. We discuss our observations in the context of performative aspects of human actions in public space, and how they can be fed back into the design of gesture interfaces for public displays.

Categories and Subject Descriptors
H.5.2 [Information Interfaces and Presentation (e.g., HCI): User Interfaces – input devices and strategies (e.g., mouse, touchscreen), interaction styles (e.g., commands, menus, forms, direct manipulation), user-centered design).

General Terms
Design, Human Factors.

Keywords
Public displays, playful behaviour, performative interactions, full-body interaction, natural user interfaces.

1. INTRODUCTION
Public displays represent a type of pervasive display that is rapidly finding its way into all public areas of the city, from TV screens in pubs to large LED screens at public squares. How these new displays can be adapted to allow passers-by to interact with their content, however, is still a topic of discussion. Naturally, the HCI community has taken on this challenge and developed a large body of work that investigates a range of methods to enable interaction with public displays (e.g. [9, 17, 18, 19, 20, 28]). For example, touch input has been promoted to support natural user interaction that can be translated to public displays [17]. However, touch input is simply impractical for large displays, such as the one we used in our study, as it was designed to be viewed from a distance of at least 3 metres, to take in its full effect. In such scenarios, the use of mid-air gestures [21] is a more promising interaction method, as they allow for interaction at a distance.

Building on previous work on gesture interfaces [10, 20, 28], we therefore developed a public display application that allowed passers-by to explore an information space using four simple gestures [1]. The application was developed for a projection-based public display that was located at the façade of a university building, facing the courtyard of an adjacent theatre (Figure 1, top). Consequently, the original brief we were tasked with was to design an application for passers-by to learn both about events at the theatre, and activities in the university. To make passers-by aware of the interactive capabilities of our display [9] and as a form of user feedback [20], our application presented people standing within the interaction range of the gesture sensor with a mirror image. We chose an abstract representation in the form of a skeleton (Figure 1, bottom) over a silhouette or live video feed, as this user representation avoided concealing the main information shown on the display. A previously published analysis of interaction logs collected over 120 days showed that our application indeed succeeded in attracting passers-by to stop and face the display [1] and that at least 2 out of the 4 gestures were easy to learn and apply [2].

However, a considerable number of our interaction logs showed long periods of people facing the screen with only a few gestures being successfully triggered. Through a preliminary analysis of the corresponding depth video recordings, we noticed a surprising level of playful behaviour, that is, an intrinsically motivated engagement with the display that did not seem to have a direct goal [13]. We attribute this playful behaviour to a number of factors in our display setup, including the novelty of the interaction [23] as well as the spatial configuration of the display and its interactive zone [3, 6, 12], amplified by the somewhat whimsical mirror representation in form of a skeleton.

In this paper we present an analysis of interactions at the display to better understand (i) the types of playful behaviour that passers-by engage in when shown their mirror representation in a public information display, and (ii) how and when such behaviour can
participants with interactive mechanisms are often cited as ‘cat has proved to be a recurring with a more focused and serious engagemen [13] situated outside of everyday life and with no direct benefit or goa [5] Playfulness to be strongly biased towards playfulness and performance. In ge [23, 27], social interaction easily emerged, with people quickly self-organising in groups for collaborative play. However, the vast majority of people made use of the interactive features in a playful and performative way (e.g. throwing photos to each other), with only a few taking a more serious look at its actual content. Such intrinsic motivation to play, to the detriment of a more serious engagement with the content, has also been observed in utilitarian applications, such as the touch-based Worlds of Information [17] or our own gesture-based public information display [9], MyPosition [27], which allowed passers-by to vote on civic topics, might be one of the few examples that succeeded in getting people to interact with the content rather than merely playfully exploring the interactive features. This was achieved by presenting people with personal traits of previous participants (in the form of silhouettes or photographs). In other words, accountability for their opinions increased the perceived seriousness of interactions with the display. Performative interaction [7, 14, 15, 25] is inherent to contexts where media technology is deployed in public spaces and where interfaces are large enough to encourage bodily interaction and allow other people to watch those interacting. Contextual constraints such as location and spatial configuration play a significant role in promoting performative interaction [3, 12] and in the emergence of the honeypot effect [4, 11, 19]. Likewise, passive engagement has been described as an essential human need fulfilled by public spaces [5] and studies have shown that, ultimately, the presence of people interacting with a display is one of the strongest factors attracting more people to the space. It is understood that an essential part of the experience of interactive environments is, to a great extent, defined by the passive act of

act as a gateway into a serious exploration of the displayed content. Based on our findings we suggest strategies for how an otherwise utilitarian information display can respond to playful behaviour, and how the identified types of playful behaviour can inform the design of gesture-based public interfaces.

2. RELATED WORK

The arrival of low-cost sensors, such as the Microsoft Kinect camera, led to a large number of HCI studies on public interactive displays, from evaluating the intuitiveness of gestures [10] to testing the effectiveness of mirror representations for attracting user interaction [20]. Typically these studies either present the user with an interface to perform a very simple task (e.g. [10, 21]) or a custom-developed game (e.g. [20, 28]). There are only few studies of gesture-based public information displays that evaluated how passers-by responded to and interacted with the display in the wild, which we discuss below.

In general, interaction with large displays in a public space seems to be strongly biased towards playfulness and performance. Playfulness refers to activities that are intrinsically motivated, situated outside of everyday life and with no direct benefit or goal [13]. Performance, in the context of our analysis, consists of a set of intentional activities by an individual or group who have some established knowledge about the interactive context and are in the presence of, and for, another individual or group [7]. Reconciling the playful and performative aspects of public interactive settings with a more focused and serious engagement with their interfaces has proved to be a recurring challenge. On the one hand, playful interactive mechanisms are often cited as ‘catalysts’ to interaction with public displays [22] as well as social interaction between participants [11]; on the other hand, they promote ephemeral participation and discourage a more in-depth engagement with the content. In CityWall [23], social interaction easily emerged, with people quickly self-organising in groups for collaborative play. However, the vast majority of people made use of the interactive features in a playful and performative way (e.g. throwing photos to each other), with only a few taking a more serious look at its actual content. Such intrinsic motivation to play, to the detriment of a more serious engagement with the content, has also been observed in utilitarian applications, such as the touch-based Worlds of Information [17] or our own gesture-based public information display [9], MyPosition [27], which allowed passers-by to vote on civic topics, might be one of the few examples that succeeded in getting people to interact with the content rather than merely playfully exploring the interactive features. This was achieved by presenting people with personal traits of previous participants (in the form of silhouettes or photographs). In other words, accountability for their opinions increased the perceived seriousness of interactions with the display.

Performative interaction [7, 14, 15, 25] is inherent to contexts where media technology is deployed in public spaces and where interfaces are large enough to encourage bodily interaction and allow other people to watch those interacting. Contextual constraints such as location and spatial configuration play a significant role in promoting performative interaction [3, 12] and in the emergence of the honeypot effect [4, 11, 19]. Likewise, passive engagement has been described as an essential human need fulfilled by public spaces [5] and studies have shown that, ultimately, the presence of people interacting with a display is one of the strongest factors attracting more people to the space. It is understood that an essential part of the experience of interactive environments is, to a great extent, defined by the passive act of
people observing others to interact in public [24]. Walter et al. [28] point out that mid-air gesture-based interfaces, encouraging a great level of bodily interaction, can be regarded as particularly performative. Active participants therefore find themselves balancing between three simultaneous processes: (1) their interaction with the public display, (2) the perception of themselves within the situation, and (3) the social role they perform for the local audience. Such split focus is inherent to the interaction with large public displays and is always shaping the user’s understanding and perception of the interaction as ‘performative spectator’ and ‘spectating performer’ [6].

Performative interaction is therefore linked to the presence of spectators, which commonly is the case in a public space. It might further occur in a highly serious context and with a clear goal in mind, while the playful behaviour we observed in our study was purely relating to ‘play’ as an activity without pursuing a specific goal [13]. Playful behaviour does not depend on the presence of spectators, however our study suggests that the performative character of a situation may actually encourage it.

Although, as this review shows, playful interaction with public displays has been observed in previous studies, the types of playful behaviours occurring at a gesture-based public display have so far not been subject to an in-depth analysis. We close this gap by identifying three types of playful behaviour from data collected over an 8-day period. Based on this analysis, we present strategies for how our findings can guide future implementations of gesture-based public display interfaces.

3. SYSTEM OVERVIEW

Media Ribbon is an application designed for a public display located at the School of IT on the main campus of the University of Sydney. The display consists of two high-performance projectors that are lined up to produce a 1.2 by 4.2 meter large display area with an aspect ratio 32 by 9. The back projection film is attached from the inside onto a glass wall facing the courtyard of a nearby theatre (Figure 1, top). A single Kinect camera is mounted in an enclosure in the center of the projection area. The interactive sweet spot is marked via yellow tape attached to the ground in the form of a cross, approximately 3 meters from the display. The display is on every day from 6pm to midnight.

The media ribbon application itself was an iteration of a previously deployed information system [1]. It features a number of items grouped into categories, represented in the form of a three-level hierarchy. The items of the currently selected level are displayed in the form of a ‘ribbon’, with the central item filling almost the entire vertical space of the display, and the items to its left and right shown in reduced size (Figure 1, bottom). For our study, the application included five different datasets: information about the nearby theatre (2 sublevels), the research clusters in the Faculty of Engineering and IT (2 sublevels), the research activities in the research group that developed the display (1 sublevel), entries from an alumni photo competition (2 sublevels), and Lastly, information about events for a light & ideas festival that was happening at the time when we performed the study (1 sublevel). Thus, there were a total of 116 items, with 1 to 11 elements per level, in the form of photos and a short text description.

The system supports a total of four gestures: left, right, more, and back. The left and right gestures are performed by moving the left or the right hand from the side of the body to the front of the body, akin to a swipe performed by moving the entire arm. The gestures are mapped to the media ribbon, to navigate left and right through the series of displayed items. The more-gesture was performed by holding the right arm up for 1 second, the back-gesture by holding the right arm downwards at a small angle for 1 second. These mappings were chosen to reflect the way the currently selected level moves up to the top edge of the screen when performing the more-gesture to delve into an item. Similarly, the back gesture brings the previous level back from the top to the centre of the display.

To help passers-by to learn the gestures available in the interactive display, the application shows four static pictograms at the bottom edge of the display. This placement follows the recommendation from Walter et al., who found that spatial separation was the most effective since it does not interrupt the main content area [27]. The pictograms further provided feedback about the execution of gestures. When a gesture was successfully completed, the corresponding pictogram was highlighted.

4. STUDY SETUP AND METHODS

We conducted the study during a winter festival of light & ideas, which took place in Sydney in May to June 2013. There were special shows as part of the festival at the theatre facing the public display. The application ran from 6pm until midnight for a total of 10 nights. We discarded the recorded data from the first 2 nights, since most of these interactions were from testing the display and due to some issues with logging, fixed on the second night.

For the remaining 8 days, we recorded the depth image video stream from the Kinect camera, the screencast from the application, as well as the number and length of interactions performed at the display. A preliminary analysis, in which we randomly looked at video recordings of varying length, revealed that playful interactions usually resulted in long engagement with the application. Recordings of shorter interactions were less likely to involve playful interactions. We therefore sorted all recorded occurrences of interactions by their length and chose the 40 longest recordings (out of 1,011) for further analysis. The duration of the recordings ranged from 59 to 397 seconds (M=126, SD=89.34). A total of 102 people featured in these recordings.

For the analysis, we watched the video depth images for each of the 40 recordings. In our analysis, we identified three distinctive types of playful behaviour: dancing, locomotion, and gesturing (described below). For each of the recordings, we captured type(s) of playful behaviour, number of people, and additional observation notes. Since the depth camera colour-coded the people appearing in each recording, we used these colours to describe the characters in a recording. For example, one of the captured comments was: “Blue attempted only 1 gesture. Blue and Red watch from the back as Purple and Yellow interact.”

5. TYPES OF PLAYFUL BEHAVIOUR

Of the selected 40 recordings, 29 featured some playful behaviour (i.e. at least one person in the recording engaged in playful behaviour). Across these 29 recordings, a total of 49 people engaged in playful behaviour (out of the 102 people across all 40 recordings). We further found that 38 of these 49 people also engaged with the content displayed on the public display by successfully performing at least one of the left, right, more and back gestures. We now focus on the 29 recordings (49 people) that featured at least one form of playful interaction.

5.1 Dancing

We use the classification ‘dancing’ to refer to repetitive and rhythmic motions. Dancing was found to be a very common form of playful behaviour. A total of 30 people engaged in some form of dancing – playing with the skeleton, possibly also triggered
through the performative nature of the interaction. The fact that the display was situated in a highly public space as well as its location and spatial configuration [3, 12], affected the interaction, with people not only being aware of themselves and the display, but also of others around them [6]. In many of the scenes, we observed one person performing with one or more people standing beside them, watching (in an expression of the honeypot effect). The dancing ranged from subtle dance moves at a fixed spot, to spinning around and moving from side to side (Figure 2). This behaviour was surprising, given that interaction with public interfaces has been observed to cause social embarrassment [4, 22]. We suspect that people saw the skeleton mirror image as legitimisation for their playful behaviour, making their dance performance (from their perspective) socially acceptable.

5.2 Gesturing
We classified playful interactions as ‘gesturing’ if they involved performing some form of hand or arm gesture other than the gestures needed to interact with the application. We observed gestures such as posing (e.g. flexing – Figure 3, bottom), expressive gestures (e.g. flapping arms – Figure 3, middle), and gestures directed at the mirror representation (e.g. shadowboxing or waving – Figure 3, top). We observed 35 people gesturing, making this the most common form of playful behaviour, observed. This is likely due to the skeleton representation; the slightly whimsical proportions of torso, head, and limbs may have encouraged people to move their arms and hands to explore how their movement changed their skeletal reflection on the display. We further attribute the popularity of this playful behaviour to the fact that gesturing is less conspicuous than some of the other behaviours we observed, therefore allowing people to playfully explore the display without risking social embarrassment.

5.3 Locomotion
Frantic movement in front of the display, such as jumping and running around, was classified as ‘locomotion’. Locomotion was the least common form of playful interaction recorded, with only 11 people engaging in this form of behaviour. Of these, 7 involved a person jumping (either on the spot, or from side to side – Figure 4, top) and 4 people running up (Figure 4, bottom) and down or shuffling on the spot. We attribute the low number of locomotion interactions to the fact that such frantic movement would be quite likely seen to be more socially embarrassing than the other observed behaviours. We further suspect that the skeleton representation prompted dancing and gesturing more than locomotion.

5.4 Discussion
Most of the people featured in the analysed scenes engaged in more than one form of playful behaviour. A common pattern seemed to be for people to transition from one playful behaviour to another, and only after that, attempt to interact with the application. In some cases however, the first person walking into the scene would attempt to interact, then transition to playing with their skeleton representation, or their nearby onlookers would step in and playfully perform. For example in one scene, the first person did some of the gestures, then two nearby onlookers also did some, but then switched to playing with their skeleton representation. This transition happened after a total of 51 seconds, suggesting that if people do not immediately see a benefit from the interaction, they will move on, finding a more compelling form of engagement.

The findings from our analysis suggest that the larger the number of people in one scene – i.e. the more performative the context – the more likely it is for playful interactions to occur. A Chi Square significance test confirmed an effect of group size on playful behaviour (p < 0.025). This observation can be linked back to previous work, which described the triangular relationships that are typically in place to trigger performative interaction [19] and the importance of spectatorship and performance to the experience of play [8]. The relationships include the awareness of the user themselves, the awareness of the interface (the display), and the awareness of onlookers. In fact, in our setup, we observed two different types of relationships between the interacting person and onlookers: the first type was their relationship with, and awareness of, nearby onlookers, which in most cases seemed to be people they knew; the second type was their awareness of distant onlookers, i.e. people near the theatre’s entrance as well as people sitting at the bar tables outside the theatre (Figure 1, top).

The skeleton representation seemed to act as trigger for playful behaviour, since it was the only element in the application that

Figure 2: Frames from one of the depth video recordings, showing two people engaging in dancing behaviour.

Figure 3: Depth image captures showing three forms of gesturing: shadowboxing (top), waving arms (middle), and posing (bottom).
visually responded to the types of playful interactions we observed. Yet, participants kept performing those interactions, even though it must have been evident that their actions did not change the state of the application.

Interestingly, we found that a number of people engaged in playful behaviour in the form of an ‘exit pattern’. After interacting with the display and actively exploring the application, they would jump into a short playful interaction as they were walking away from the display. The most common exit patterns we observed were people dancing as they were walking away or waving at their skeleton reflection in the display.

6. IMPLICATIONS FOR DESIGNING GESTURE-BASED INTERFACES

Although we designed our application as a serious interface, we derived a number of unexpected insights from the analysis of playful behaviour observed at the display, which we describe in this section. We acknowledge that our findings are limited in that we analysed a selected group of 40 out of 1,011 recordings logged over a period of 8 days. Our insights are further strongly linked to the display’s location and spatial configuration, the application design, and the use of a skeleton for representing passers-by. In our future work, we plan to compare the effect of different user representations, e.g. skeleton versus silhouette, on the level of playful interactions from passers-by.

6.1 From playing to interacting

Our analysis revealed that 38 of the 49 people demonstrating playful interactions also engaged with the content displayed on the public display. This means that they successfully performed at least one of the left, right, more or back gestures. To ensure these were not accidentally triggered interactions, we compared the depth video recording with automatically logged screen captures of the Media Ribbon interface. Out of those 38 interactions, a total of 15 people started with playfully engaging with the display and subsequently moved on to interacting with the application. We attribute this pattern to two observations: first, the public nature and exposure of the person’s actions, making them aware and more conscious about their actions [27]; and second, the catalyst effect of the skeleton reflection for drawing people’s attention to the display [9], and then consequently discovering the supported gestures for exploring the content. This suggests that allowing people to play with the display may be an effective strategy for drawing people into an interaction in an otherwise serious context.

6.2 Supporting spontaneous gestures

We observed a number of gestures that people seemed to naturally perform, most likely triggered by their mirror representation in the display. In our application, these gestures did not have any effect, since the set of gestures was limited to four navigational gestures for exploring the hierarchical information in the application. We believe that the usability of gesture interfaces, which by their very nature are difficult to learn [28], could be improved by supporting those naturally occurring playful gestures. This is supported by our observation of 23 people, who seemed to attempt the provided gestures at first, but then after a short period transitioned to playing with the display. For example, a number of people attempted gestures that were not supported, such as raising their arm from the thigh to make a 90-degree angle with the torso, or waving with their hand at the display. Rather than ignoring unsupported gestures, the application could display a contextually appropriate response, such as a help message in response to someone waving at the display. In a few instances we also observed people performing collaborative gestures, e.g. by holding hands, and observing their reflection in the display. The application could respond to such gestures e.g. by unlocking special features [29] or bringing up a multi-user mode.

6.3 Adaptive gesture interfaces

A challenge with designing public interactive displays is the fact that it is hard to predict the demographics, motivations, and objectives of the people interacting with the display [6, 22]. In fact, these parameters might change depending on the time of the day and other circumstances [26]. For example, many of the people interacting with the display were on their way to see a show at the theatre, or more often, coming out of the theatre after having seen a show. The demographics were therefore linked to what was showing each particular night. For instance, on the first day of our study period, there was an event hosted by the university, attracting a more academic audience, whereas the fifth and sixth days (Friday and Saturday) saw a music line-up, with a total of nine local bands performing over the two days. Based on our analysis of playful behaviour, we see an opportunity to design public display applications as responsive, fluid interfaces that adapt automatically to the crowd based on the playfulness of the interactions. For example, if people walk up to the display in large groups and start engaging in highly expressive gesturing behaviour, the application could automatically change to a mode in which the content is presented in a more playful form. Similarly, the available gesture vocabulary could be adapted, e.g. to support full-body interaction, which has been successfully applied in public game interfaces [20, 28].

7. CONCLUSION

Without intending to provoke playful behaviour, we observed a large number of unprompted playful interactions at our public display application. We attribute this to the display’s performative context and spatial configuration [3, 12], the novelty of the interaction [6], and the somewhat whimsical representation of passers-by in the form of a mirrored skeleton. Our analysis of 40 recordings of interactions revealed that only 35 out of 102 people featuring in these recordings interacted with the application without showing any form of playful behaviour. Almost half of them (49 out of 102) engaged in some form of playful behaviour.
Based on our analysis we presented insights for the design of public interactive displays. Firstly, we observed that a number of people transitioned from playing with the display to exploring the information shown on the display. We therefore suggest that carefully using playful elements in public displays may help to attract interaction, as well as potentially aiding the learning of mid-air gestures. Secondly, we identified a number of gestures, such as waving and the ‘exit pattern’, and suggested ways for using these in public display applications. Thirdly, we proposed that the type and level of playful behaviour could be used to automatically adapt the public display interface to better suit the demographic and objectives of passers-by.

8. ACKNOWLEDGMENTS

This work is partially funded by the Smart Services Corporate Research Centre (CRC) as part of the Multi-Channel Content Delivery and Mobile Personalisation Project.

9. REFERENCES


