Measuring Interactivity at an Interactive Public Information Display

Christopher Ackad\(^1\), Rainer Wasinger\(^1\), Richard Gluga\(^1\), Judy Kay\(^1\), Martin Tomitsch\(^2\)

1. Faculty of Engineering and Information Technologies; 2. Faculty of Architecture, Design, and Planning
The University of Sydney, NSW, 2006, Australia
{christopher.ackad, rainer.wasinger, richard.gluga, judy.kay, martin.tomitsch}@sydney.edu.au

ABSTRACT
Public Information Displays (PIDs) have only recently begun to support user interaction. Traditionally, such displays have been static and non-interactive, and past research has shown that users of such displays (both non-interactive and interactive) are often oblivious to them; a term commonly known as ‘display blindness’.

In this paper, we describe the results from a field study that was conducted on a gesture-based PID, to observe interactivity with the display over a number of different experiment conditions. Over a period of 120 days, a total of 2,468 people approached the display. Results show that 71% proceeded to face the display, and from this, 62% of these people proceeded to interact with the display, with average interaction sessions lasting 28 seconds. Results from this study provide valuable insight into interaction sessions with interactive PIDs, as well as an essential baseline for future studies into PID interactivity.

Author Keywords
Interactive public information displays, gestural interaction, user centered design and user studies, ubiquitous computing

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation: User Interfaces

INTRODUCTION
A number of technological advancements have recently enabled Public Information Displays (PIDs) to become interactive. Among these technological advancements are the availability of new vision- and gesture-based technologies (e.g. Microsoft’s Kinect\(^4\) and the Leap Motion controller\(^5\)), improvements to the accuracy of speech recognition engines (e.g. Google Voice\(^3\)), the use of smartphones as a communication device (e.g. for pointing and navigating such interfaces \([4, 2, 8]\))

Experimentation into how and why people use interactive PIDs continues however to go largely unexplored, despite many challenges currently existing with these public displays. Some of these challenges stem from the traditional non-interactive nature of the displays, which have in the past been often employed for the sole purpose of advertising. One such challenge is known as ‘display blindness’, and occurs when users - who expect uninteresting display content - ignore the display entirely \([7]\).

A more recent challenge for interactive PIDs is how to effectively and efficiently convey the display’s interactivity to passers-by. In \([3]\), a number of different visual interface conditions were evaluated to help increase the percentage of people who face a display. The authors in \([6]\) also perform studies in this area, with a particular focus on display noticeability and the effect that the use of a silhouette has on increasing user interaction compared to a more traditional attract sequence with visual call-to-action such as a banner and the text ‘Step Close to Play’. Following in the footsteps of the audience funnel paradigm \([5]\), the challenge that this work addresses is how to inform a user of the vocabulary with which to interact with an interactive display. This is particularly relevant for interfaces based on gesture and/or speech, as these modes of interaction are by their very nature invisible, making it challenging to communicate the functional boundaries of the application to the user. Compare this for example to a GUI, in which graphical elements like buttons and menus are visible and persistent to the user at all times.

THE INTERACTIVE PID INSTALLATION
In this paper, we describe the results of an extensive field study that was conducted on an interactive PID installation located in close vicinity to the Seymour Centre theatre and the School of IT building at the University of Sydney. The goal of this study is to observe user interactivity with the display across two separate experiment conditions. The first experiment condition tests interactivity of advertising. One such challenge is known as ‘display blindness’, and occurs when users - who expect uninteresting display content - ignore the display entirely [7].

A more recent challenge for interactive PIDs is how to effectively and efficiently convey the display’s interactivity to passers-by. In [3], a number of different visual interface conditions were evaluated to help increase the percentage of people who face a display. The authors in [6] also perform studies in this area, with a particular focus on display noticeability and the effect that the use of a silhouette has on increasing user interaction compared to a more traditional attract sequence with visual call-to-action such as a banner and the text ‘Step Close to Play’. Following in the footsteps of the audience funnel paradigm [5], the challenge that this work addresses is how to inform a user of the vocabulary with which to interact with an interactive display. This is particularly relevant for interfaces based on gesture and/or speech, as these modes of interaction are by their very nature invisible, making it challenging to communicate the functional boundaries of the application to the user. Compare this for example to a GUI, in which graphical elements like buttons and menus are visible and persistent to the user at all times.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

\(^{1}\)Microsoft, URL: www.microsoft.com/kinecret/orwindows/
\(^{2}\)Leap Motion, URL: www.leapmotion.com
\(^{3}\)Google Voice, URL: www.google.com/googlevoice/
\(^{4}\)MicroTouch ClearTek, URL: www.3m.com

Copyright 2013 ACM 978-1-XXXX-XXXX-XX/XX...$10.00.
on the facility’s research clusters). The second experiment condition focuses on interaction with the display in a tutorial-mode and a self-discovery mode. Over the course of 120 days (60 days for each experiment condition), a total of 2,468 gestures were recorded. The results of this study, as outlined later in this paper, provide valuable insight into how users interact with interactive PIDs.

Placement and Hardware
Our interactive PID installation has been deployed at the end of a glass building that runs along a pedestrian walkway and is accessible from an adjacent theatre courtyard. This area experiences mid-range pedestrian traffic as well as a peak in stationary traffic during events in the courtyard and prior to shows that run in the adjacent theatre. The installation consists of two high-intensity 1080p projectors, projecting on a rear projection film that has been laminated onto the glass wall. A single Kinect for Windows hangs in its own enclosure above and in front of the projection display as seen in Figure 1.

User Interface
The main UI element is the ‘Media Ribbon’, a horizontal scrolling set of media items. These media items can be text, images, or videos (Figure 3, Figure 4). The intention is to provide a simple and intuitive user interface to allow people to quickly navigate and view content. Below this we have a real time representation of the user’s skeleton rendered on top of the displayed content and a help bar that displays icons representing the four available gestures to the user (Figure 2).

Based on previous experience [3], we chose to implement four upper body gestures (‘Swipe Left’, ‘Swipe Right’, ‘More’ and ‘Back’), each quite distinct, in order to simplify interaction and improve reliability. ‘Swipe Left’ and ‘Swipe Right’ are used to navigate right and left in the UI. ‘More’ and ‘Back’ are used to drill down into the sub-menu or to display more information and to navigate back to the previous ribbon. The ‘More’ gesture is triggered when a person holds their arm up, and the ‘Back’ gesture is triggered when they hold their arm down at a 45 degree angle from their waist, similar to the Xbox 360 pause menu gesture.

STUDY DESIGN
As outlined in [1], commercial public displays are often used for advertising (e.g. in airports, railway stations, and high streets), digital signage solutions (e.g. in public buildings), and for making locations more attractive to users (e.g. screens in bars and cafes). Such content is however not typically designed for public displays and thus does not cater well to the properties of digital public displays and the short interaction times associated with such displays. The authors in [1] further outline how although display owners are experimenting with interactivity, the majority are still reluctant to deploy interactive displays.

The goal of the study outlined in this paper is to see how successful our public information display is at attracting passers-by to view and engage with the display. This is measured using characteristics of user interaction sessions with interactive public information displays (i.e. interaction conversion rates, average interaction durations, and content consumption rates) to determine the level of success. Another goal is to determine how such interaction can be affected by the introduction of a tutorial compared to a static display and the type of content shown. We measured interaction with this public displays in terms of the number of people who have passed by the display (referred to in this work as ‘skeletal arrivals’), faced the display (i.e. ‘facing display’), and have
interacted with the display (i.e. ‘interacting users’). One similar study to which this work can be compared is [6], in which the details of a pre-study are outlined in which passers-by interact with a game-based interactive display that allows users to play with interactive balls. In that study, 832 passers-by were observed, 54.8% of whom faced the display and from those that faced a further 37.5% interacted with the display.

The field study outlined in this paper was conducted over a period of 120 days from February to May. During this period, two separate experiment conditions were tested. The first focused on content type (i.e. theatre content versus research facility content) and the second focused on tutorial type (i.e. non-tutorial versus tutorial). The experiment was conducted such that the interactive PID showed the theatre content set for the first 60 days, and the research facility’s content for the second 60 days. During each of these 60 day periods, the experiment was further divided by the tutorial conditions, where each condition (tutorial and non-tutorial) was run for a total of 30 days. The system was programmed to automatically start at 18:00 and shutdown at midnight each day (a total of 720 hours over the 120 day period). This evening time slot was particularly well suited to the theatre patrons attending evening performances.

Differences between the two content sets were as follows:

- **Theatre content**: The theatre content (Figure 3) was a single level content set that had 24 items, all on performances to be shown at the theatre.

- **Research clusters content**: The research clusters content (Figure 4) was, by contrast, a multi-level content set of 30 items, divided into 7 top-level items, each of these with a further 3-5 sub-level items.

Similarly, the differences between the tutorial types can be described as follows:

- **Non-tutorial**: In the non-tutorial condition (Figure 3), people who were detected by the system (i.e. skeletal arrivals) were provided a welcome message at the top of the display (‘Welcome! Use your hands to interact with the display’), as well as the set of four gestural icons, which were displayed at the bottom of the screen along with the skeletal representation of the user. To help alert the user to the presence of these gestural cues, the interface would cycle through each of the four grey icons and highlight them in yellow at a rate of 3 seconds per icon.

- **Tutorial**: In the tutorial condition, skeletal arrivals were also provided a welcome message at the top of the display, though this time the message was accompanied with text requesting the user to try the first of the four available gestures (‘SWIPE your LEFT hand to navigate LEFT’). As shown in Figure 4, the tutorial consisted of four ordered tasks, each one representing a separate gesture: ‘left’, ‘right’, ‘more’, and ‘back’. In this mode, each current gestural icon in the tutorial was highlighted in white and the completion of each gesture would see the icon turn yellow. On completion, the user interface would return to the non-tutorial state.

**RESULTS**

During the 120 day period, a total of 2,468 skeletal arrivals were detected by our interactive PID. From this number, a total of 1,746 users faced the display (i.e. 70.75%) and from those that faced, a further 1,078 users (or 61.74%) interacted with the display. In comparison to the related work outlined earlier [6], these numbers indicate a substantial conversion rate both from those passing-by to facing, and again from those users facing to interacting with the display. One possible reason for the high interaction rates (compared to past work) is that our recordings are based on the skeletal arrivals detected by the Kinect sensor rather than a human observer; the Kinect sensor has a maximum detection range of 5m, which means that people passing further than 5m away from the display are not counted in our analysis. Another possible reason may just be that the content displayed was sufficiently interesting for passers-by to interact with it. Also interesting to note is that although our system had high conversion rates between the different phases of the audience funnel [5], the resulting number of content items viewed was still relatively low; on average 2.3 items (10%) of the theatre content and 2 items (7%) of the research clusters content.

The results in Table 1 show that users interacted more in the non-tutorial mode over the tutorial mode, with consistently higher proportions for those facing and interacting, as well as higher numbers for the content items viewed and the average interaction durations. This is counter-intuitive, as we would have expected this tutorial mode to encourage people to interact with the display more. One possible cause for this discrepancy is thought to be that passers-by are generally not willing to invest much time (e.g. to complete a tutorial) before being able to interact with the display. This is further supported in Table 2, which shows the decrease, or funnel effect, in the use of the four gestures. The ‘left’ and ‘right’ gestures occur at similar rates, with a range of 63% to 74% of users across the conditions performing one of them. There is a large drop in moving to the ‘more’ gesture, which was used by only 14% for the research content where it was essential for exploring the hierarchy, and 33% for the theatre content where this made the selected item much larger and showed additional information. This drop is similar across the tutorial on and off modes (22% and 25%).

**CONCLUSIONS**

This paper has presented an interactive public information display and the results of an extensive field study that provide valuable insight into how users interact with a gesture-based interactive PID. This work is particularly relevant to interactive public displays that support natural user interactions based on gestures, because...
<table>
<thead>
<tr>
<th>Condition</th>
<th>Arrivals</th>
<th>Facing Display</th>
<th>Interacting Users</th>
<th>Content Viewed (items)</th>
<th>Interaction Time (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tutorial - Off</td>
<td>1045</td>
<td>761 (73%)</td>
<td>497 (65%)</td>
<td>2.43</td>
<td>29.4</td>
</tr>
<tr>
<td>Tutorial - On</td>
<td>1423</td>
<td>985 (69%)</td>
<td>581 (59%)</td>
<td>1.96</td>
<td>27.1</td>
</tr>
<tr>
<td>Content - Theatre</td>
<td>1209</td>
<td>792 (66%)</td>
<td>468 (59%)</td>
<td>2.32</td>
<td>30.1</td>
</tr>
<tr>
<td>Content - Research</td>
<td>1259</td>
<td>954 (76%)</td>
<td>610 (64%)</td>
<td>2.07</td>
<td>26.4</td>
</tr>
</tbody>
</table>

Table 1. Tabulation of users that faced and interacted with the interactive PID during the experimental study.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Left</th>
<th>Right</th>
<th>More</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tutorial - Off</td>
<td>328 (66.10%)</td>
<td>355 (70.47%)</td>
<td>105 (24.57%)</td>
<td>64 (14.34%)</td>
</tr>
<tr>
<td>Tutorial - On</td>
<td>403 (69.35%)</td>
<td>389 (66.92%)</td>
<td>129 (22.24%)</td>
<td>74 (12.75%)</td>
</tr>
<tr>
<td>Content - Theatre</td>
<td>312 (66.63%)</td>
<td>293 (63.45%)</td>
<td>152 (33.31%)</td>
<td>82 (17.91%)</td>
</tr>
<tr>
<td>Content - Research</td>
<td>419 (68.82%)</td>
<td>451 (73.94%)</td>
<td>82 (13.50%)</td>
<td>56 (9.18%)</td>
</tr>
</tbody>
</table>

Table 2. Tabulation of the different gesture types performed by users during their interaction with the interactive PID.

These interfaces often require a learning phase that may be a barrier for people passing by a public installation.

The results of our 120 day field study show that compared to past work, this interactive PID has a high conversion rate of passers-by that first faced the display and then that transitioned from facing to interacting with the display. This shows that passers-by are willing to be engaged by interactive PIDs. Our results further show that interaction with PIDs need to be intuitive and that a tutorial mode can be counter productive, particularly in cases where passers-by are not willing to spend much time (e.g. 30 seconds or less) interacting with the display.

ACKNOWLEDGMENTS
This work is partially funded by the Smart Services CRC as part of the Multi-Channel Content Delivery and Mobile Personalisation project.

REFERENCES


