Increasing Passersby Engagement with Public Large Interactive Displays: A Study of Proxemics and Conation

Mojgan Ghere¹, Marvin Palla¹,², Caroline Wong³, James R. Wallace⁴, Stacey D. Scott¹²

¹Department of Systems Design Engineering  
²School of Computer Science  
³Department of Management Sciences  
⁴School of Public Health and Health Systems  
University of Waterloo  
Waterloo, Ontario, Canada  
{mojgan.ghare|mpafla|ck4wong|james.wallace}@uwaterloo.ca

ABSTRACT

Prior research has shown that large interactive displays deployed in public spaces are often underutilized, or even unnoticed, phenomena connected to ‘interaction’ and ‘display blindness’, respectively. To better understand how designers can mitigate these issues, we conducted a field experiment that compared how different visual cues impacted engagement with a public display. The deployed interfaces were designed to progressively reveal more information about the display and entice interaction through the use of visual content designed to evoke direct or indirect conation (the mental faculty related to purpose or will to perform an action), and different animation triggers (random or proxemic). Our results show that random triggers were more effective than proxemic triggers at overcoming display and interaction blindness. Our study of conation – the first we are aware of – found that “conceptual” visuals designed to evoke indirect conation were also useful in attracting people’s attention.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords
interaction blindness; display blindness; public interaction; large displays; multi-touch; proxemics; conation

INTRODUCTION

Public large interactive displays (PLIDs) are rapidly being adopted for public interactions due to their ability to deliver dynamic content to a broad audience, to support interactions by groups of individuals, and their increasing affordability. Recently, these devices have even begun to replace traditional paper signage in some public areas [38], across contexts as broad as interactive photo displays [44], semi-public whiteboards [46], and kiosks [3, 34].

Researchers have found that PLIDs often receive limited usage, or even go unnoticed. While individuals are largely familiar with and eager to use personal devices like tablets and smartphones, they often do not recognize PLIDs in the field, or overlook their interaction capabilities [38]. These challenges have been formalized in the literature as ‘display blindness’ [41] and ‘interaction blindness’ [25, 42], respectively, and are identified by a growing number of studies that suggest these ‘blindnesses’ are key barriers to deploying PLIDs in practice. While formalizing these issues is an important first step, those wishing to deploy PLIDs lack guidance on how they can be designed to engage the general public.

To better understand how designers can overcome display and interaction blindness, we conducted an in-the-wild field study that compared the usage patterns of different PLID interface designs. In particular, we investigated two types of visual cues that are thought to influence the use of public displays and signage: proxemics (e.g., [4, 21]) and conation (e.g., [41, 25]) and conation [35]. Inspired by sociologist Edwin T. Hall’s [21] theories on the use of space, proxemic displays actively respond to users based on their proximity to the display [19], and have been shown to engage people from a distance (e.g., [8, 54]). Conation, or conative thought, is one of three aspects of mental processing, and is defined as the mental faculty of impulse, purpose or will to perform an action [13]. We were inspired by the use of conation-based design in advertising, and explored the use of direct and indirect conation to encourage passerby engagement with PLIDs.
Our analyses of more than 2600 encounters over a 4-day period contributes insights into how visual cues can be leveraged to encourage interaction with PLIDs in a real-world setting. First, our results show that our proxemic-aware interface was less effective than non-proxemic interfaces in overcoming interaction blindness with our PLID. Second, our analysis of conation points to its potential as a tool to mitigate display blindness, rather than interaction blindness as the advertising literature would suggest [34]. We describe these results, and discuss their implications for the design of PLIDs.

RELATED WORK
In the Human-Computer Interaction (HCI) research community, Public Large Interactive Displays (PLIDs) have been envisioned as a platform for interaction for decades (e.g., [16, 33, 55]). Over time, that vision moved from supporting closed-door meetings with a small group of individuals, to public displays available for anyone to walk up and use. For example, Vogel and Balakrishnan proposed 'Public Ambient Displays' [54] that provide passersby with more detailed and private information as they become close to the display. As technologies improved, displays have gained resolution and high-powered graphics capabilities, more accurate and responsive multi-touch capabilities and body-tracking; and have become larger — but the goal of supporting spontaneous and serendipitous interactions by passersby has remained the same.

However, as these technologies have been deployed to the field, researchers have often observed that PLIDs are not always understood, or even noticed, by the general public. Passersby may choose to not interact with a PLID for a number of reasons — for example, many may experience social embarrassment [6, 15, 18], be afraid of breaking the device [42], or simply have low expectations or motivation towards public display content [41]. But, research has pointed towards passersby being 'blind' to a PLID's functionality. In particular, two types of blindness have been identified as design challenges by the HCI community:

Display blindness occurs when passersby fail to notice the presence of a PLID in a public setting, even when it is in close proximity [26, 41]. Display blindness can apply equally to both interactive (i.e., PLIDs) and non-interactive (e.g., paper) displays, and has been observed extensively by HCI researchers in practice (e.g. [22, 28]).

Interaction blindness occurs when a display is seen by someone, but is is not recognized as interactive [25, 42]. Overcoming both display and interaction blindness is also referred to as the 'first click problem' [27, 28, 34].

We now review related literature that seeks to address these two types of blindness.

Display Blindness: Drawing Passersby’s Attention
Overcoming display blindness requires designers of a PLID to understand how to entice passersby to ‘notice’ a display [8]. However, HCI research has found that this is particularly challenging, and that designers may have a very narrow window of opportunity to influence passersby. For instance, in a field study that tracked participants’ gaze while walking through a shopping centre, Dalton et al. [14] found that passersby often only very briefly looked at displays. They report that participants would often look at PLIDs for approximately 1/3 of a second and that 96% of glances lasted less than 800ms.

To draw a passerby’s attention to a PLID in this narrow window of opportunity, the research community has often explored the use of ‘low level’ visual stimuli [9] that humans can process quickly, with little higher-order cognitive processing. These stimuli include visual channels such as colour, luminance, and animation (motion). Dalton et al. [14] reports changes in luminance were largely ignored by their participants, and may not be a useful design tool for attention. Similarly, the use of colour is often context- and content-dependant, and may not always be available to designers. However, motion is more readily used to capture passersby’s attention, i.e. to overcome display blindness, and research has explored different means of doing so (e.g., [8, 28, 38]). In this work, we focus on two approaches to the use of motion in PLIDs: proxemics and animated interface elements.

The term ‘proxemics’ was coined by Hall [21], and is based on interpersonal distances (proximity) often observed in the field between humans, comprising, for example, ‘personal’, ‘social’, and ‘public’ spaces. In the context of large displays, these proxemic zones have been leveraged to make PLIDs ‘react’ to passersby in a more socially conscious way. For example, Brudy et al. [7] and Zhou et al. [57] have explored the use of proxemics to protect privacy while interacting on multi-touch displays. In the context of display blindness, our earlier work [9], which built on work by Ballendat et al. [4] and Müller et al. [39, 40], conveyed motion through a passersby’s silhouette displayed on the PLID. The visual design of the silhouette grew larger and more salient as the user approached the display, drawing passersby’s attention and encouraging them to explore the PLID further.

Animation has also been applied more directly to interface elements in the display. Kukka et al. [28] compared the ability of animated icons and text to attract users to approach and touch a PLID. They report that animation was one of the most successful ways of capturing passersby’s attention, but in discussing their results note that previous work has reported the opposite effect [11, 51]. Of particular interest is that Kukka et al. report that PLIDs with textual interfaces attracted passersby more often with animation, whereas graphical (i.e., icons) PLIDs did not. We discuss these potential design choices in more depth when describing our experimental platform.

Interaction Blindness: Touching the Display
In contrast with low level stimuli, high level stimuli are more complex visual elements that require cognitive processing to interpret and understand (e.g., on-screen text or icons) and, thus, are more likely to help persuade passersby to interact with the PLID [29, 38]. Consequently, the research community has largely explored how textual and graphical (i.e., icons) interfaces influence engagement of passersby.

Notably, Kukka et al. [28] investigated the ‘first click problem’ [27, 34] through a field deployment of 8 displays that were assigned either textual or iconic interfaces. They found that
interfaces featuring text elicited more interaction than those featuring icons. However, in reflecting on their results, they explain a limitation of their work: “Regarding text and icon, it is challenging to make a reliable comparison – it is probable that the selection of words and the design of icon impact the effectiveness of such signals” (p. 1706).

Lösch et al. [30] explored how different spatial arrangements of displays foster engagement. In particular, they found that adding nearby displays that showed life-sized video representations of passersby increased the number of people who stopped to interact, but decreased the number who interacted with the primary display. In this work, we focus on single display designs, and alternative metaphors for engaging passersby.

In seeking to address interaction blindness, we were inspired by research by McQuarrie and Phillips [35], and more generally from the advertising literature to explore how different means of conveying a message might impact engagement with PLIDs. McQuarrie and Phillips [35] argue that indirect, even suggestive, images can often be more convincing than text. They explain “that indirect claims, such as those using metaphor, may be advantageous because they render the consumer more receptive to multiple, distinct, positive inferences” (p. 7). Their research relies largely upon the premise that alongside affect and cognitive thought, humans also rely upon conative thought to make decisions. Conative thought is simply defined as one’s impulse, and determines how one acts on affect and cognitive thoughts [2]. In this work, we explore how conation can be used to overcome interaction blindness with PLIDs. In particular, we investigate whether passersby may be receptive to suggestive images.

**EXPERIMENTAL PLATFORM**

The literature suggests that design for display and interaction blindness must address two different levels of visual stimuli: ‘low level’ stimuli can prompt passersby to notice a display, whereas ‘high level’ stimuli play a large role in encouraging those passersby to decide to interact with it. Our review identified promising stimuli: use of motion through animation or proxemic interfaces to address display blindness, and conative design to overcome interaction blindness.

To better understand how these design choices may ultimately impact a passerby’s interactions with a PLID, we developed an experimental platform to enable us to conduct controlled experiments in a field setting, and to measure passersby’s interactions with a deployed PLID.

We decided to implement a community information kiosk based on past findings by Müller et al. [41] that found passersby often ignore public displays due to expectations of advertising rather than useful information. To overcome this perception, and to suit our anticipated student audience on a university campus, we developed a campus community large-display application that provided access to useful information such as bus schedules, availability of nearby study space, and the length of line-ups at nearby coffee stores.

Our design uses a bookshelf metaphor (Figure 2), where different types of information are presented as books. This design was chosen based on a number of useful properties: it allows new ‘books’ to be added as new content becomes available, is familiar to the intended audience, and also allowed for us to incorporate other objects that can respond to passersby to capture their attention. We designed a background image for the bookshelf on which these objects can be overlayed (e.g., the toy bicycle and lamp in Figure 3). Further, we designed the bookshelf to be at eye height for the average passerby, with the middle shelf displaying primary content and the upper and lower shelves reserved for these awareness objects.

The platform was developed in Unity 5, with a passerby’s proximity data captured by a Microsoft Kinect V2. As identified in the discussion of related literature, our platform focused on the investigation of two independent visual variables: INTERACTIVE and CONATIVE mode. Items placed on the shelf were designed to appear and react to passersby differently, depending on which modes were active. We now describe how each of these modes worked.

**Interactive Modes**

Our platform enabled items on the bookshelf to respond to passersby using one of two INTERACTIVE modes: PROXEMIC and ANIMATED.

**Proxemic Mode**. Under the PROXEMIC mode, items respond to passersby as they enter one of three ‘zones’ in front of the display: 0 - 1m, 1 - 1.5m, and further than 1.5m (Figure 3). This three-zone approach was previously adopted in our prior PLID work [9, 8], and is similar to the ‘personal’, ‘subtle’, and ‘social’ spaces in Vogel and Balakrishnan [54] and Hall’s ‘intimate’, ‘personal’, and ‘social’ zones [20]. We simply label these Zones 1, 2, and 3, in order from closest to farthest from the PLID (Figure 3).

As passersby approached the display, they would proceed through Zone 3, then Zone 2, and finally Zone 1. To react to approaching users, and encourage them to move closer to the PLID, on-screen items reacted to their position and movement. In Zone 3, awareness items like the bicycle on the top shelf of the display would roll back and forth, or books on the shelf would tilt to match a passerby’s movement. In Zone 2, books close to the user would open up, revealing text and icons that would suggest the availability of more content should the passerby decide to approach the display. Finally, in Zone 1, the display would roll back and forth, or books on the shelf would tilt to match a passerby’s movement. In Zone 2, books close to the user would open up, revealing text and icons.
To develop comparable interfaces for both DIRECT and INDIRECT conation, we iteratively designed a set of text descriptors and icons for campus services, including health services, the campus physical activity centre (i.e., the gym), nearby bus stops, coffee shops, study spaces, and campus night life. In Zone 3, the DIRECT interface simply overlayed the text descriptor over the spine of the corresponding book, whereas the INDIRECT interface overlayed a more suggestive icon. In Zone 2, when the book opens up, a more detailed icon was overlayed on the book’s pages, with different icons developed for each Conative Mode. Finally, in Zone 1, an iconic hand is displayed for the DIRECT interface, with an outstretched finger touching the display and a highlighted touch point. For the INDIRECT interface, a magnifying glass is shown to encourage passersby to investigate the interface.

Importantly, while both DIRECT and INDIRECT conation may feature text or icons, there are subtle differences in their approach to engaging users. The DIRECT conation approach is designed to be like a traditional user interface that announces functionality in the most literal way possible through textual labels that describe functionality or representative icons, e.g., the words ‘Discover more...’ accompanied by an icon of a hand touching the display or a photograph of a cafe. Our design approach for INDIRECT conation is motivated by advertisement design and makes use of more suggestive icons and text, such as the need to recharge to suggest coffee shop locations, or an icon of a magnifying glass to suggest that more information is available.

A challenge in designing DIRECT/INDIRECT interfaces is that the difference between the two is not always clear cut, and similar interface elements may be used in different stages of interaction. For example we use the text ‘Coffee Shop’ in both interfaces, but in different zones. For DIRECT conation, it is used to provide more context about the types of information available on the display in Zone 3, and to entice passersby to interact [8], whereas it appears in Zones 1 and 2 for INDIRECT conation. Lack of prior work in the literature meant that we needed to do our best to embody the two design philosophies. We tried to do what made sense as passersby progressed through each Zone, and to maintain consistency within each conative design philosophy.

FIELD EXPERIMENT
To assess the effectiveness of the Interactive and Conative Mode designs, we deployed our PLID to a hallway in an engineering building at the University of Waterloo campus (Figures 1 and 4). The location is adjacent to an above-ground pedestrian pathway that is frequently used by students between classes. It is also near staff offices and classrooms, and thus staff members also commute through the area between meetings or classes. We deployed the PLID during the Canadian Winter term, a few weeks before midterm reading week. Hence, we also expected that cold, seasonal weather and students’ preparation for the midterm exams would increase traffic through the area.

Similar to past studies of PLIDs [8, 28, 44], our goal was to capture passersby’s interactions in a naturalistic setting through both quantitative and qualitative measures. By deploying the display over a number of days, we were able to capture interactions with a large and diverse set of passersby, compared to a laboratory environment.
Study Design
Our 2 × 2 between-subjects field experiment included two independent variables: **Conative Mode** and **Interactive Mode**, corresponding to the designs included in our experimental PLID platform. **Conative Mode** was either Direct or Indirect, with on-display designs consisting of text or icons, respectively. **Interactive Mode** was either Proxemic or Animated, with the display either responding to participant’s proximity to the display, or randomly cycling through a series of animations.

As we were interested in the impact on these design approaches to addressing both display and interaction blindness, our study measures include both behavioural and subjective measures. Passersby’s behaviour was captured during the study and video coded during post-experiment analysis, as explained below. This video coding focused on identifying observable behaviour that indicated passersby’s attention to the display (e.g. looking and stopping) as indications of whether a given experimental condition was successful at countering display blindness. Observable behaviour that indicated passersby’s interactions with the system (either attempts at interaction, such as trying mid-air gestures in front of the display, or effective interaction such as touching the display) were taken as a measure of a condition’s effectiveness at countering interaction blindness. Survey data were also collected from a subset of passersby as a self-report measure of what design aspects captured their attention or invoked (or discouraged) their interaction.

Experimental Setup
Our experimental software was installed on a SMART Board 6000 series touch interactive 1920 × 1080 65” display, mounted on a mobile stand of adjustable height. The display was connected to a Windows 8 PC (3.5GHz CPU, 16GB RAM, NVIDIA Quadro K2200). A Microsoft Kinect V2 was mounted to the top of the display to enable tracking users. The Kinect sensor was situated in a black cardboard frame atop the display to minimize its presence, as our prior experiences have found that passersby often react to a visible Kinect and form expectations about what type of user input the system expects (e.g. mid-air gestures) based on their prior experiences with or knowledge of the consumer Kinect device.

Procedure
The PLID was deployed for four days in mid-February, from 11am - 3pm each day. The order of presentation of **Conation** and **Interactive Modes** was counter-balanced across the four days, using a partial Latin-square design, with each mode displayed for 1 hour. While the PLID was deployed, we observed unobtrusively and recorded field notes regarding any interesting passersby behaviour. Once a passersby had successfully interacted with the PLID, and determined that they should seek out the “red ribbon”, an investigator verbally recruited them to participate in an interview. Moving participants away from the display reduced confounds in our experiment due to the honey pot effect, and allowed investigators to speak to participants in more detail about why they approached the display.

Interviewees were first read a verbal recruitment script, and if interested, the investigator asked them to read the cover page of electronic survey as well as the information consent. The interviewee confirmed their agreement verbally and then responded to a set of predefined questions via electronic survey on a tablet. They were also asked a series of verbal questions, and responses to those questions were recorded by the investigator. The study protocol was approved by our institutional research ethics board. This approval included the stipulation that a sign be posted in the study location the week following the field study with information on the study and the contact information of the researchers and the research ethics office should anyone have any questions or concerns with the study.

Data Collection and Analysis
To better understand the impact of proxemics and conation on gaining passersby’s attention and engagement (i.e. overcoming display and interactivity blindness), we collected both quantitative and qualitative data related to passersby’s behaviour and perceptions. Passersby’s positional data within the PLID’s three Zones and their touch interactions on the display were captured in computer log files. Interactions with the PLID were also video recorded with a Sony HDR-MV1 handheld video camcorder. Passersby who agreed to participate in the on-site survey answered questions about their initial impressions of the PLID and its visual content. The survey consisted of closed/open-ended survey questions.

To determine the extent to which passersby attended to and interacted with the display, we conducted an analysis of the collected video data. We followed a closed coding approach [48], using an a priori coding scheme adapted from our prior PLID work [8] that included the following behavioural codes:

- **None** Shows no intention to look or pause with display in view, walks by as if it is not there.
- **Glanced** The action of looking at the display without stopping. A head turn towards the display is equal to a glance.
- **Stopped** The action of stopping and looking at the display.
- **Explored** Anyone who showed active exploratory behaviour in front of the display, such as trying to determine whether the display supported mid-air gestures, or the general purpose of the system.
- **Touched** The action of touching the display.

A group of people or an individual at the display were each treated as a single unit of analysis. Thus, for individuals we only coded the first occurrence of a given behaviour; if a passersby touched the display multiple times, we recorded only one occurrence. Similarly for groups, we only coded the first target behaviour observed from any member of the group. For instance, if someone from the group looked at the display, this group would be coded as “Glanced”, but if a second group member looked at the display their behaviour was not coded. Consistent with prior public display research, we consider the “first” occurrence of a given behaviour in a group context to be most likely due to the display, rather than influenced by others in the group. Through this video analysis, we identified each
passersby that glanced at, stared at, explored, and touched the display, and to what experimental conditions they were exposed.

To understand the efficacy of the Interactive and Conative design approaches, we conducted a logistic regression analysis on coded ‘Touch’ behaviour. This statistical test was chosen due to the binary nature of our dependent variable (i.e. touch or no touch), and its ability to consider multiple independent variables, look for interaction effects, and, most importantly, to predict the relationship between our independent and dependent variables [17]. The logistical regression provides a simple, but statistically robust measure of the impact each design approach had on engaging passersby.

Similar to prior research [36, 31, 43, 56], we then conducted a conversion analysis to understand the impact of each design approach on a passersby’s progress through the “audience funnel”, as coined by Michels and Müller [37]. That is, the narrowing band of engagement from passing by the PLID, to stopping, to interacting with the PLID. This analysis examines the efficacy of a design approach at eliciting target behaviours along the funnel from passersby, including Glanced, Stopped, Interacted (any form), and/or Touched. Note, Interacted (any form) combines Explored and Touched behaviors.

The conversion analysis comprised a series of comparison tests on the frequency of observed target behaviours, one for each stage of the funnel. For example, for the Glanced stage, we compare the number of passersby who Glanced to the number who did not (e.g., Not Glanced). For successive tests along the funnel, the data were filtered by the pool of passersby who performed the prior behaviour in the funnel. For example, for the Explored stage, only those people who Stopped were considered, as only they would have had the opportunity to explore or touch the display. This analysis helps us to identify whether a certain design approach can foster people’s transition to the next stage. For instance, a certain design may attract many passersby to stop, but be less effective at enticing interaction.

To understand the impact of Interactive and Conative Mode on passersby’s progress through the funnel, we first conducted a conversion analysis that collapsed the data into these two factors, running separate conversion analyses for each. For these analyses, Fisher’s Exact test was used because of its robustness to small cell counts (e.g., Touched behaviour for some conditions). We then conducted a conversion analysis across each of the four interfaces. We also used Fisher’s Exact test for this analysis, followed by post-hoc pairwise comparisons, with Bonferroni corrections. All statistical tests were performed with \( \alpha = .05 \).

RESULTS

Over the course of the 4-day period of the study, 2613 passersby (groups and individuals) were coded within the vicinity of the display. Of those passersby (all coded as Passed by), 853 of them Glanced at (32.64%), 97 Stopped at (3.71%), 59 Explored (2.26%), and 51 Touched (1.95%) the display. We first report the logistic regression for Touched behaviour. Then we report our conversion analyses for the experimental factors (Interactive and Conative Modes) and the individual conditions (each separate interface), which incorporates all target behaviours to better understand the impact of our design approaches on the entire interaction progress. Finally, we report our qualitative findings, including general observations and our survey results.

Logistic Regression Analysis on Touching Behaviour

The logistic regression analysis results are summarized in Table 1. The analysis found a significant main effect for the Interactive Mode (beta coefficient= -.902, \( p = .0037 \)), but not the Conative Mode (beta coefficient=0.345, \( p = .2314 \), n.s.). The regression model was setup with Animated and Direct Conation as the baseline levels for each factor. Thus, the negative coefficient for Interactive Mode indicates that switching from Animated to Proxemic decreased the odds of passersby touching the display, or that significantly fewer passersby in the Proxemic condition touched the display than those in the Animated condition. No significant interaction was found between the factors, so this term was omitted from the model. In general, our model is significant \( \chi^2(2) = 11.32 \), but only explains a little variance in the dependent variable (Nagelkerke’s \( R^2 = .025 \)).

Conversion Analyses

Table 2 summarizes the results of our conversion analyses for Interactive and Conative Modes, and for the study conditions. The table includes the frequency of target behaviours observed throughout the study period, as well as the results from the Fisher’s Exact test for each analysis.

Conversion Analysis for Interactive and Conative Modes

The conversion analysis of Interactive Mode found that passersby were significantly less likely to Glance at \( (p < .001, \text{odds ratio}= .74) \), Stop at \( (p = .007, \text{odds ratio}= .54) \), or Touch \( (p = .019, \text{odds ratio}= .14) \) the display in the Proxemic conditions than in the Animated conditions. No significant difference was found for Interacted (any form) behaviour. This analysis shows that the Animated design approach facilitated progression through the audience funnel more effectively than the Proxemic approach, effectively attracting people to the display as well as communicating its interactivity. Note, while not significant, the frequency data reported in Table 2 show that very few passersby who Interacted in some way in the Animated conditions did not Touch the display (5%), whereas many more passersby (29%) in the Proxemic conditions Interacted but did not Touch (i.e., they performed alternative exploration behaviours such as mid-air gestures). We discuss this observation later.

The conversion analysis of Conative Mode found that passersby were significantly more likely to Glance at the display \( (p < .001, \text{odds ratio}= .74) \) in the Indirect Conation conditions than in the Direct Conation conditions. A marginally significant difference was found for Stopping behaviour \( (p = .067, \text{odds ratio}= .51) \), with passersby more likely to stop at the display in the Indirect Conation conditions. No significant differences were found for any of the other target behaviours. This analysis shows that Indirect Conative conditions were better at attracting attention to
Table 1. Summary of the logistic regression testing the impact of the experimental factors on Touched behavior.

<table>
<thead>
<tr>
<th></th>
<th>$b_i$</th>
<th>$se_i$</th>
<th>Odds ratio and CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-3.741**</td>
<td>0.2437</td>
<td>0.0143 0.0237 0.0371</td>
</tr>
<tr>
<td>INTERACTIVE MODE</td>
<td>-0.902**</td>
<td>0.3107</td>
<td>0.2143 0.4058 0.7313</td>
</tr>
<tr>
<td>CONATIVE MODE</td>
<td>+0.3452</td>
<td>0.2884</td>
<td>0.8067 1.4122 2.5167</td>
</tr>
</tbody>
</table>

Notes: $R^2 = 0.023$ (Hosmer-Lemeshow), $0.004$ (Cox-Snell), $0.025$ (Nagelkerke). Model $\chi^2(2) = 11.32$, **$p < .01$, ***$p < .001$. 

Table 2. Summary of the conversion analysis for INTERACTIVE MODE (top), CONATIVE MODE (middle), and the four separate conditions (bottom). The frequency of observed target behaviors are provided, along with Fisher Exact test results for the different target behaviors considered in each conversion analysis.

<table>
<thead>
<tr>
<th>INTERACTIVE MODE</th>
<th>Not glanced</th>
<th>Glanced</th>
<th>Not stopped</th>
<th>Stopped</th>
<th>Not explored</th>
<th>Explored</th>
<th>Not touched</th>
<th>Touched</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANIMATED</td>
<td>821</td>
<td>462</td>
<td>397</td>
<td>65</td>
<td>27</td>
<td>38</td>
<td>2</td>
<td>36</td>
</tr>
<tr>
<td>PROXEMIC</td>
<td>939</td>
<td>391</td>
<td>359</td>
<td>32</td>
<td>11</td>
<td>21</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>Fisher-Test</td>
<td>$p &lt; 0.001$</td>
<td>$p = 0.007$</td>
<td>$p = 0.52$</td>
<td>$p = 0.019$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Odds Ratio</td>
<td>0.74</td>
<td>0.54</td>
<td>1.35</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95% CI</td>
<td>[0.63, 0.87]</td>
<td>[0.34, 0.87]</td>
<td>[0.52, 3.65]</td>
<td>[0.01, 0.92]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CONATIVE MODE</th>
<th>Not glanced</th>
<th>Glanced</th>
<th>Not stopped</th>
<th>Stopped</th>
<th>Not explored</th>
<th>Explored</th>
<th>Not touched</th>
<th>Touched</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIRECT</td>
<td>926</td>
<td>411</td>
<td>373</td>
<td>38</td>
<td>12</td>
<td>26</td>
<td>5</td>
<td>21</td>
</tr>
<tr>
<td>INDIRECT</td>
<td>834</td>
<td>442</td>
<td>383</td>
<td>59</td>
<td>26</td>
<td>33</td>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td>Fisher-Test</td>
<td>$p = 0.037$</td>
<td>$p = 0.067$</td>
<td>$p = 0.287$</td>
<td>$p = 0.284$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Odds Ratio</td>
<td>1.19</td>
<td>1.51</td>
<td>0.59</td>
<td>2.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95% CI</td>
<td>[1.01, 1.41]</td>
<td>[0.96, 2.40]</td>
<td>[0.23, 1.49]</td>
<td>[0.41, 16.8]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

By Condition  Not glanced | Glanced | Not stopped | Stopped | Not explored | Explored | Not touched | Touched |
| ANIMATED/DIRECT          | 396       | 209     | 188        | 21       | 6            | 15       | 1           | 14      |
| ANIMATED/INDIRECT        | 425       | 253     | 209        | 44       | 21           | 23       | 1           | 22      |
| PROXEMIC/DIRECT          | 530       | 202     | 185        | 17       | 6            | 11       | 4           | 7       |
| PROXEMIC/INDIRECT        | 409       | 189     | 174        | 15       | 5            | 10       | 2           | 8       |
| Fisher-Test¹             | $p < 0.001$ | $p = 0.006$ | $p = 0.463$ | $p = 0.050$ |

¹ The Fisher Exact test does not provide odds ratio with more than two variables.

the display than the DIRECT CONATIVE conditions, but neither design approaches provided a significant advantage for communicating interactivity.

Conversion Analysis of Individual Conditions
The conversion analysis of the individual conditions found a significant difference across conditions for the Glanced ($p < .001$) and Stopped ($p = .006$) behaviours, and a marginally significant difference for Touched behaviour ($p = .050$).

For Glanced, pairwise comparisons found that passersby were less likely to Glance at the PROXEMIC/DIRECT CONATION condition compared to either the ANIMATED/DIRECT CONATION ($p < 0.05$, odds ratio = 0.72, 95% CI = [0.57, 0.92]) or ANIMATED/INDIRECT CONATION ($p < 0.001$, odds ratio = 0.64, 95% CI = [0.51, 0.81]) conditions.

For Stopped, pairwise comparisons found that passersby were more likely to stop at the display in ANIMATED/INDIRECT CONATION than in either the PROXEMIC/DIRECT CONATION ($p < 0.05$, odds ratio = 0.44, 95% CI = [0.23, 0.81]) or PROXEMIC/INDIRECT CONATION ($p < 0.05$, odds ratio = 0.41, 95% CI = [0.20, 0.78]) conditions.

For Touched, pairwise comparisons found no significant differences after family-wise error adjustments.

This analysis shows that the ANIMATED conditions were more effective than the PROXEMIC conditions at attracting a passerby’s attention and motivating them to stop and investigate the display. ANIMATED/INDIRECT CONATION had the strongest effect for glancing and stopping behaviour.

Qualitative Results
Over the four days of deployment, we made a number of observations about how passersby engaged with the PLID, and surveyed a number of those who ultimately touched the display. We observed that many passersby appeared to be familiar with the concept of a display system that tracks your movement and actions (like we were doing for the PROXEMIC conditions), in particular, the Microsoft Kinect hardware. Some passersby
even peeked behind the display presumably to investigate the hardware being used. Although we anticipated this curiosity, and disguised the Kinect within a black cardboard frame, passersby tried various body and hand gestures in front of the display (Figure 4). Many of these passersby were thus classified as Exploring the display, even though they may not have explored whether the PLID was touch-sensitive.

Consistent with many prior PLID studies, we observed a strong ‘honeypot effect’ [40], where passersby approached the PLID after seeing others using it. In many cases, through group exploration, an individual leader would experiment with different body and hand gestures while others observed from a distance. However, when groups of only two people approached the display, they would either touch the display simultaneously or invite each other and explore possible feedback together.

Out of those passersby that touched the PLID, 35 (25 male, 9 female, 1 other) participated in our survey. Participants were primarily in the 18-24 age group (33 participants), and 89% had previously used a PLID in a public setting. We asked participants ‘What initially drew your attention to the display?’ (Table 3). Responses largely reflected the (large) physical display, its colourful interface, and other passersby already interacting with it. ‘Other’ responses included novelty of the display (2), that it ‘looked cool’ (1), and even ‘[I] designed some of the PCBs in this [physical display] model’ (1).

We also asked participants ‘What encouraged you to approach the display?’ (Table 4). The most frequently reported responses were, ‘seeing others at the display’ (37%), ‘curious about information’ (29%), and ‘moving interface objects’ (14%). Finally, we asked participants ‘What interface elements made you think the display was, or might be, interactive?’ Participants most frequently said ‘button-like objects’ 18 (51%), followed by ‘moving objects’ 15 (43%), and 1 participant responded ‘books opened when I walked by’.

**DISCUSSION AND IMPLICATIONS FOR DESIGN**

Our results provide insight into design choices that can influence passersby’s attraction to and engagement with PLIDs. They point to the potential of indirect conation for addressing display blindness and benefits of randomly triggered animations – and drawbacks of proxemics-triggered animations – for overcoming display and interaction blindness. We now reflect on our findings, how they compared to our expectations before running the study, and their implications for PLIDs.

**Table 3. Responses to ‘What initially drew your attention to the display?’**

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>The physical display device</td>
<td>18 (51%)</td>
</tr>
<tr>
<td>The colourfulness of visual content</td>
<td>14 (40%)</td>
</tr>
<tr>
<td>‘I saw someone using the display’</td>
<td>11 (31%)</td>
</tr>
<tr>
<td>Moving objects in the UI</td>
<td>7 (20%)</td>
</tr>
<tr>
<td>The ‘Community’ title</td>
<td>5 (14%)</td>
</tr>
<tr>
<td>Other</td>
<td>4 (11%)</td>
</tr>
<tr>
<td>My familiarity with the visual content</td>
<td>2 (6%)</td>
</tr>
</tbody>
</table>

**Table 4. Responses to ‘What encouraged you to approach the display?’**

<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seeing other people use the display</td>
<td>13 (37%)</td>
</tr>
<tr>
<td>Curious about information I could get</td>
<td>10 (29%)</td>
</tr>
<tr>
<td>Moving interface objects</td>
<td>5 (14%)</td>
</tr>
<tr>
<td>Curious about possible changes in the interface</td>
<td>5 (14%)</td>
</tr>
<tr>
<td>Previous experience with PLIDs</td>
<td>1 (3%)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (3%)</td>
</tr>
</tbody>
</table>

**Increasing the Attraction Power of PLIDs**

The attraction power of a PLID – a concept borrowed from museum studies to measure the ability of an exhibit to attract visitors [5, 24] – denotes its ability to both attract passersby’s attention and entice them to stop and investigate the display. The first trait strongly relates to a display’s ability to overcome display blindness, while the second one relates to the display’s ability to intrigue and/or communicate utility to passersby. Attracting a passerby is the first step toward overcoming the ‘first click problem’ [27, 28, 34].

The base interface design of our experimental platform – the community bookshelf – was intentionally designed to be attractive to our expected population, university students and staff. Its educational library theme with content related to campus activities was meant to appeal to the interests of this population, while its bright, colourful visual design was meant to visually appeal to passersby. Indeed, even in the “worst” performing condition (PROXEMIC/DIRECT CONATION), the display was noticed by 38% of passersby, and in the “best” condition (ANIMATED/INDIRECT CONATION), it was noticed by almost 60% of passersby. These results are consistent with reported Glanced statistics in prior studies, including 6-16.2% [26], 28.8-41% [47], and 23-60% [14], across various public contexts. Yet, our results also show that both changes in animation triggers and iconography impacted the attraction power of our PLID.

**Proxemic- vs. Randomly-Triggered Animations**

Prior research suggests that proxemics-based animation triggers, such as changing the speed of moving images [9] or displaying silhouettes that mirror a passersby’s movements [40], can increase the attraction power of a PLID. We drew

---

1. Note, our coded Glanced behavior was limited to visibly observable head turns toward the display, and does not capture instances where passersby glanced with their eyes without turning their head.
2. Dalton et al. [14] recorded “glances” for displays with more than 10 passersby.
The tendency for proxemics-based visual cues to be missed as people approach a PLID, before they are in the range of cues designed to encourage interaction helped provide more attraction power, as randomly generated sequences of visual cues may be helpful for mitigating display blindness, it still limits movement. However, our results suggest, while this approach the potential visibility of visual cues that mirror a passerby's cues slightly in front of the passersby's position to increase have previously suggested, is to project the mirrored graphical surfaces as recently explored by Lösch et al. [30]. An alternative mitigation strategy, which we [10] and Schmidt et al. [49] have previously suggested, is to project the mirrored graphical cues slightly in front of the passersby's position to increase the potential visibility of visual cues that mirror a passerby's movement. However, our results suggest, while this approach may be helpful for mitigating display blindness, it still limits attraction power, as randomly generated sequences of visual cues designed to encourage interaction helped provide more insight into the overall possibilities of the display and likely encouraged people to stop and investigate further. Moreover, as people approach a PLID, before they are in the range of body tracking, randomly generated visual cues are already visible and can help attract attention and interest.

Indirect vs. Direct Conation-based Iconography

Given our design rationale for use of indirect conation, these results may at first seem counter-intuitive. That is, we included the indirect conation design approach in our study with the expectation that it would evoke ‘high-level’, more cognitive engagement with the display — and encourage interaction with the display, as opposed to addressing display blindness. However, one explanation for the impact on display blindness is that graphical elements in the indirect conation interface tended to have more saturated colours when compared to the text (in Zone 3) and photographs (in Zone 2) used in direct conation interface. This explanation is supported by our qualitative data, in which passersby who responded to our survey cited ‘the colourfulness of visual content’ as the second-most frequently mentioned reason that the display drew their attention.

While the indirect conation design approach did not prove to be as effective as we had hoped at persuading passersby to visit the display or interact with it (as discussed further below), our data did show a trend (with marginally significant differences, \( p = .07 \)) for the indirect conation conditions to result in more stopping behavior. However, reflecting upon our results, we feel that indirect conation may not be effective for the complexity of PLID interfaces typically explored by the surface computing research community. That is, interfaces with many different informational items and interaction possibilities. The advertising literature tends to feature much simpler graphics with one or two major components, that enable the reader to more carefully focus and engage with that content. Our implementation may have featured too many graphical elements, overwhelming passersby with stimuli, and inhibiting their ability to engage with content on the deeper level required for conative messaging to be effective.

We believe that indirect conative design may be more effective when the interface is trying to convey one or two simple messages, or in contexts where interactive surfaces are more focused on a specific task. The first approach, was actually incorporated into our base interface design – the community bookshelf. This interface provided a conceptual representation of information — a library — rather than the more typical interface buttons to represent information that could be obtained from the display. An example of the second approach would be incorporating indirect conation into the design of interactive mannequins that support shopping for clothing [45].

The design choices made in developing our platform are also only one example of a conative interface, and different choices may well lead to more effective interfaces. This is the first study that we are aware of that specifically investigated conative design, and addresses a gap identified by Kukka et al. [28] in trying to understand how different textual and iconic design patterns may influence use of PLIDs. Having engaged in this study, we find that we agree with Kukka et al.’s conclusion that measuring the effectiveness of any one design element is challenging, and that there is a need for further study and for triangulation with more in-depth laboratory studies (and
Consistent with prior PLID studies, our results found that 1.0-3.2% of our 2613 passersby touched the display. Sample meaningful” engagement levels reported in other studies include 1.5-1.7% [12], 2.2-2.9% [43], 1.4-5.6% [32], and 3.4-5% [47]. Together, these results show how difficult it is to engage passersby in meaningful interaction. Engaging every passerby is not the goal of public display design; not everyone will be interested or have time to engage with the display. Rather, the goal is to remove any barriers for passerby who might be interested or benefit from display engagement. To this end, our results show that the ANIMATED interfaces better achieved this goal than our PROXEMIC interfaces did. In particular, we found that the ANIMATED interfaces had over 2 times as many passersby touch the display than the PROXEMIC interfaces. These findings are consistent with those of Kukka et al. [28], who also found that animations are an effective way of enticing passersby to interact with a PLID, but provide an additional, novel comparison to proxemic interfaces (e.g. [4, 8, 54]), which have been noted as a promising paradigm by the HCI community for some time. These findings suggest a need to give pause, and to consider the implications of our findings on the use of proxemic interfaces in practice.

As we noted in our review of related literature, proxemics is often explored as a novel, intuitive, and simple means of interaction with large displays, particularly in public space. However, our results suggest that it may not be particularly effective in drawing in passersby to begin with, and that randomly triggered animations were more effective in these cases.

One explanation for these differences is that our implementation limited proxemic interactions to items on the virtual shelf, which may not be perceived by passersby as indicating an interactive display. For instance, on-screen avatars, such as shadows or skeletons, are particularly effective at fostering engagement in practice [1, 8, 39, 53]. We observed 38 passersby who noticed the Kinect device mounted above the display and ‘explored’ its interface with hand gestures while the ANIMATED interface was active. While this observation may initially be interpreted as passersby misunderstanding the display, 36 of those 38 ended up touching the display. We interpret these findings as indicative of the passersby already determining that the display was interactive, based on their observations of the display and its on-screen animations, and instead were exploring which modalities were supported.

Another explanation is social anxiety — passersby may have decided not to approach the PLID once they realized it was reacting to their movements to avoid any chance for social embarrassment in a public setting [6, 42]. Interestingly, this behaviour may protect against the risk of misusing proxemic interfaces to create a captive audience as discussed by Greenberg et al. [19] in their research on ‘Dark Patterns’. Our results are encouraging in that they suggest that dark patterns may be less useful to abuse in practice, but also challenge the notion that proxemics may be a panacea for public interaction.

Overall, these findings demonstrate that proxemics can be an effective means of interaction, but that it also has drawbacks when enticing interaction. This limitation of proxemics should be considered when deploying novel systems to the field.

LIMITATIONS
While our study was able to provide insight into how display and interaction blindness impact passersby’s engagement with public displays, some limitations should be addressed through future work. For example, the Microsoft Kinect we used to measure encounters with our PLID was not able to discern between large groups of users, requiring a secondary video analysis to confirm numbers of engaged passersby. Our choice to utilize an experimental platform in which the PLID’s functionality was ultimately limited to displaying a short message to contact a nearby investigator, may have influenced some bystanders’ choice to interact with the display and limited the ‘honeypot’ effect’s impact on interaction and display blindness [40].

Finally, as Kukka et al. [28] noted in their own discussion of limitations, the choice of specific words and icons is likely to have a significant impact on the effectiveness of a PLID, and are a limitation of any single study. We designed our experimental platform to reflect the different conative designs as best as we could, but differences between DIRECT and INDIRECT conation are often subtle and we are unaware of any existing software interfaces from which to draw inspiration. Based on our initial results, we suggest that future work should refine how conative design can be best implemented on PLIDs.

CONCLUSION
Our work contributes a deeper understanding of how proxemics and conation can be used to foster engagement with large interactive displays in public contexts. First, to our knowledge, it provides the first investigation of PLID interfaces that incorporate indirect conation to persuade passersby to stop and engage with the display. Second, our field study shows that proxemic interfaces were less effective than animated interfaces at engaging passersby, contradicting findings from the literature [4, 8, 54]. These results point to means of overcoming interaction blindness with PLIDs in public.

Moreover, our investigation corroborates a number of designs previously explored in the literature. Although the study did not find a strong main effect of CONATIVE mode on interaction blindness, the triangulation of our video coded observations, questionnaire responses, and qualitative observations provides evidence that indirect conation interfaces may help passersby effectively overcome display blindness. These findings suggest an interesting potential for interaction between proxemic and conative design that warrants further study.

ACKNOWLEDGEMENTS
We thank the Natural Sciences and Engineering Research Council of Canada for supporting this research.
REFERENCES


