"It's Mine, Don't Touch!": Interactions at a Large Multi-Touch Display in a City Centre

Peter Peltonen, Esko Kurvinen, Antti Salovaara, Giulio Jacucci, Tommi Ilmonen, John Evans, Antti Oulasvirta, Petri Saarikko
Helsinki Institute for Information Technology (HIIT)
Helsinki University of Technology and University of Helsinki
P.O. Box 9800, 02015 HUT, Finland
firstname.lastname@hiit.fi

ABSTRACT
We present data from detailed observations of CityWall, a large multi-touch display installed in a central location in Helsinki, Finland. During eight days of installation, 1199 persons interacted with the system in various social configurations. Videos of these encounters were examined qualitatively as well as quantitatively based on human coding of events. The data convey phenomena that arise uniquely in public use: crowding, massively parallel interaction, teamwork, games, negotiations of transitions and handovers, conflict management, gestures and overt remarks to co-present people, and "marking" the display for others. We analyze how public availability is *achieved* through social learning and negotiation, why interaction becomes performative and, finally, how the display restructures the public space. The multi-touch feature, gesture-based interaction, and the physical display size contributed differentially to these uses. Our findings on the social organization of the use of public displays can be useful for designing such systems for urban environments.

Author Keywords
Situated public displays, urban environments, multi-user interfaces

ACM Classification Keywords
H.5.1. Multimedia Information Systems: Evaluation/Methodology.

INTRODUCTION

Recent technological developments have made large multi-touch screens less expensive and enabled their real-world deployments. In parallel, the topic has gained more attention in the HCI community. A central feature is how the size and the "public availability" [16] of the tangible interface support simultaneous participation of multiple users. The whole surface and objects are simultaneously available through the continuity of the tangible interface and the size invites and accommodates several users.

Public outdoor spaces in urban environments are an interesting but under-researched setting for large multi-touch screens. Reported user studies have been carried out in laboratory settings or in semi-public indoor contexts such as offices and conferences. Social organization of the public space is, however, different from private settings (such as companies) and semi-public settings (such as conferences). These three settings differ in a number of dimensions that may or may not affect interaction, such as number of potential users, social relationships between potential users, and knowledge about the display and its use. To understand how multi-touch screens can affect and support social interaction, particularly in public spaces, we suggest that observational studies in urban environments should be undertaken and developed.

One of the reasons for the lack of outdoors studies has been that standard multi-touch solutions do not work well in uncontrolled lighting environments. We developed a solution that works in changing lighting conditions and created a permanent installation in a central location in the centre of Helsinki, Finland. The installation involves a multi-touch screen on a shop front window. The system, called CityWall, provides a zoomable timeline that can be panned chronologically to organize public images of the city. Pictures, downloaded from Flickr, can be resized, rotated, and moved with simple one- or two-handed gestures. CityWall serves as an open platform for passers-by to play around and explore images.

By reporting our observations of social interaction around the CityWall, we want to provide first insights into how users approach, participate, and interact on a multi-touch display in public space. In our analysis, we have treated the display as an intervention in the normal daily course of actions in the space. We are interested in the different uses of the display by passers-by, its public availability for anyone, and the differences in use during different times of day, as well as weekends, workdays and public events.

CHI 2008, April 5–10, 2008, Florence, Italy.

RELATED WORK

Much of the research that is relevant for understanding interactions between people around large interactive displays has already been carried out. This field of research is vast and varied, ranging from indoor to outdoor settings and from office and work contexts to deployments for public spaces and communities. In the following review, we focus not on the technologies as such but on findings that stem from user studies, especially ones with an interest in social uses of large displays. To improve clarity, we have categorized public interactive displays according to the social configurations and interaction they afford: tabletop, ambient and wall displays.

Tabletop displays support a particular kind of bodily configuration – standing and sitting – and are used mainly in collaborative work spaces. Research has presented new kinds of collaborative touch-based interaction techniques that take into consideration multi-hand manipulations and touching possibilities (fingers, hand angles, user-user gestures etc.) [15,18,26].[1]

Ambient displays, on the other hand, do not usually involve direct interaction on their surface. They have been developed to investigate the ways in which displays can be situated in physical settings, representing rhythms and movements of people in a space and increasing reflection and awareness of other users of space [22,24,25]. TexTales, a photograph installation in an area of buildings under a threat of being demolished attempted to develop practices of citizen journalism. An analysis of content showed eight different categories of texting, but did not address direct face-to-face citizen interaction facilitated by the display [2].

Studies on *large multi-user wall displays* are the ones closest to the case we present. Research into these displays dates back to the 1980s but only recently have studies focusing particularly on interaction been reported. The settings range from collaborative workspaces in office environments[2] to more public settings such as schools.

The study on BlueBoard, a touch-screen display that can identify users with a RFID reader, focused on possibilities to use large displays for small-group collaborative use such as sketching. This observation-based study highlighted benefits of visible physical actions (that facilitate learning from others), difficulties in developing clear turn-taking practices, and varying emerging ways to collaborate without anyone taking a leader role [19].

MERBoards were used in NASA's Mars Exploration Rover missions and studied in real collaborative settings (a control centre). In longitudinal use, the role and the function of the display changed as the use of other displays in the control room evolved [8]. However, the roles of MERBoards in team's activities were supplemental rather than central; this posed challenges related to perceived ease of use, usefulness and availability at the right time to their adoption [9].

The systems mentioned above are examples of *semi-public displays* – systems for "members of a small, co-located group within a confined physical space, and not general passersby" [7, p. 51]. By being located in indoor spaces with limited access, the content and interaction with these displays has been made fit to suit the particular work practices in that space. In contrast, *public displays* are for anyone to interact in a walk-up-and-use [10] manner. In public displays, a large proportion of users are passers-by and thus first-time users. Most of the research on public displays has been carried out by running installations in local communities. In Opinionizer [3], social interactions were studied in two social settings (a book launch party and a welcoming party for students). Two personal thresholds were found that potential users have to overcome before they can start interacting with the display. First, they have to withdraw from other activities (e.g. talking to other people). Second, once standing at the display, they still have to wait for their turn and feel willing to use the display in the presence of others. As an implication, the authors suggested positioning the display along the thoroughfares of traffic and improving the ways in which the interaction principles of the display are communicated implicitly to bystanders.

Campiello was a system installed in a local school in Venice, designed to support the local community that lives under the pressure of a high level of tourism. In this task, it was found important to gather and share collective memories and provide them to all community members in a personalized manner [1]. Another related study was about Dynamo, a multi-user system installed also in a school and designed to support multimedia content sharing. In addition to using public content, Dynamo supported the use of private content through "carving", i.e., through reservation of a dedicated space on the screen for personal purposes. During a two-week user study, use patterns evolved as users developed ways to attract other people's attention through "upsizing" their pictures, stage video performances in the display, and engage in turn-taking in relation to the amount of space collocated users could take from each other [4].

The review shows how studies of large multi-touch displays have mostly focused on tabletops and controlled settings as office spaces. We contribute to this line of work with a ethnographic observation of social uses of a large, public, multi-touch display. We elaborate earlier observations on multiple user situations, including aspects of teamwork and parallel uses, as well as interactions between strangers.

Second, earlier research has indicated turn-taking as crucial for successful collaborative use. We take a detailed look at the practices and challenges related to turn-taking, e.g. con-

[1] For a commercially available product, see, for example, Microsoft Surface at www.microsoft.com/surface/

[2] For example, Accenture's interactive wall technology, www.accenture.com/Global/Services/Accenture_Technology_Labs/Services/SeeingTheBigPicture.htm

flicts that are due because of multiple users using the screen at once. We also look at how people recognize these conflicts and how they go about managing them.

Third, we want to extend current discussion from large display prototypes to their relationship with the urban environment (see [5] for an example of how this has been explored with a different kind of technology using a portable digital carpet equipped with LED lights). Interactive installations like these can potentially restructure the way people experience and use the space around them.

INTERVENTION STUDY IN THE URBAN ENVIRONMENT

We have developed a touch screen called CityWall to investigate the interaction and situatedness of displays in an urban setting. To our knowledge, this is the first multi-user multi-touch large display that has been installed outdoors in a public space for anyone to interact with and which does not require continuous supervision or maintenance.

CityWall

The main features of the Citywall technology are 1) multiple hand tracking, capable of identifying uniquely as many fingers and hands as can fit in the screen, 2) hand posture and gesture tracking, 3) high-resolution and high-frequency camera processing of up to 60 frames per second, and 4) computer-vision-based tracking that works in changing light conditions. The main challenge was to support interactions for any user, from a child to a senior citizen, not requiring special skills or previous knowledge. The four technological features create the conditions for such a multi-user and multi-touch installation that is appropriate for public space. The set up is similar to HoloWall [14], using a rear-projection panel that is semi-opaque and diffusive, a video camera with an optical IR filter for recognition, and IR lights to illuminate objects in front of the camera directly. This setup allows us to place all the equipment indoors, and so out of the public space, and use a normal safety glass as a screen.

CityWall is especially suitable for navigation of media in general, and of photos in particular. The current version gathers content tagged with certain keywords ("Helsinki", in our case) in realtime from Flickr. Figure 1 shows a screenshot from CityWall with Flickr content displayed in it, organized according to the overview+detail principle. The bottom part (B) of the screen has a timeline with pictures in a thumbnail size. It is navigated by scrubbing it left or right and it can also be compressed or expanded to show the contents retrieved during a full day or just during a couple of minutes. This has been found important as the frequency of media may vary greatly.

Interaction with the top part (B) of CityWall follows two interaction paradigms. Moving, scaling and rotation of content (C) follows *direct manipulation* principles: a user can grab an image by putting a hand on it. The photo follows the hand movements when the user shifts her hand. Rotation and scaling are possible by grabbing the photo at more

Figure 1. Screenshot of CityWall with Flickr content.

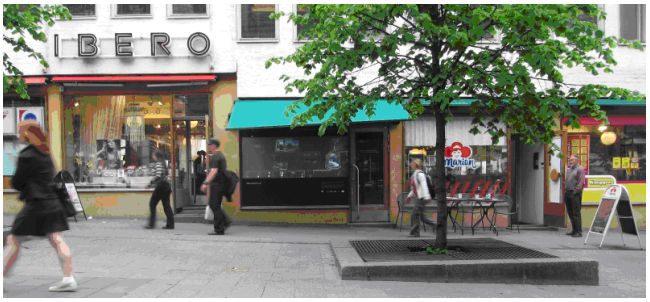

Figure 2. CityWall installation in Helsinki, Finland.

than two points (e.g. by two hands or two fingers of the same hand) and then either rotating the two points around each other or altering their distance.

The other interaction principle is *non-modality*. All the functionalities mentioned above are available for the user all the time. This is in contrast to modal user interfaces, in which different modes of interaction are often chosen from palettes or menus. Non-modality is especially important for multi-user systems because confusions arise easily if the system needs to associate different touches with different interaction modes. With non-modal interaction, this problem does not occur.

Data collection

To study touch-screen interaction in a real setting, we had an opportunity to install CityWall in a central location in Helsinki, Finland for the duration of summer 2007. The site was a 2.5-meter-wide shop window next to a café located between the main bus and train stations (see Figure 2). The two stations are used by 400 000 passengers each day, and there is a great deal of pedestrian traffic past the display around the clock.

CityWall use was recorded in multiple ways during the installation. The system wrote a continuous *interaction log* of the touches, updates of content etc. so as to timestamp the moments of interaction and see what photos were interacted with. A *web camera* was installed in the sunshade above the shop window, looking down to the street and the users. The recording was on continuously for one month (July 2007). The video captured had a 640x480 resolution and included

a 128kbps mono soundtrack. Twelve *on-site interviews* were conducted to collect immediate user feedback.

Data analysis

The video data was used as the primary content for analysis of interaction. It was first pre-processed with the help of interaction logs to leave out the video clips containing no active use. The eight last days of July were chosen for a more detailed analysis because this would provide the opportunity to see both first-time and returning users in the video. The video footage was partitioned into *sessions* of interaction, each session containing a full episode of uninterrupted use, either by one or more users. This partitioning preserved the possibility to analyze multi-user situations in which users may enter and leave the display at different times. If there was more than a ten-second gap between user interaction with the display or with each other (nobody using the wall), then the use was counted as a new session.

For each session, *manual coding* was carried out to find out (1) its duration, (2) the number of active users who touch the display, and (3) the number of passive bystanders who were using the active users. For determining the numbers of passive users, seeing their reflection in the window was helpful, as it allowed seeing a large part of the area behind the camera as well.

The data from this stage of coding were used for statistical investigations of use. Upon noticing the prevalence of multi-user interactions (see below), this data was used for two types of further analysis. The multi-user instances were subjected to a *second stage of coding* to show what group sizes were present at the display in these sessions. Because of the tediousness of analyzing whether users belonged to the same or different groups, only Monday, Tuesday and Saturday of the previously coded data were re-analyzed.

The outcome from the first coding was also used for identifying the most interesting multi-user sessions for a *qualitative analysis*. This analysis focused on the unfolding of events and interactions at the wall and the ways in which users displayed their understanding to others in these events. Material for this analysis was selected by the following selection criteria: a) the five sessions having the most active users b) the five sessions having the most passive users c) the five sessions having the most people present in total. These sessions contained sessions from different days of the week and different hours of the day.

FINDINGS

In the presentation of findings, we focus on how people used the CityWall installation, and how they collaborated and interacted with each other at the screen. The presentation draws from the statistics distilled from the coded episodes of interactions on the videotape, on-site interviews, and interaction analyses of selected multi-user episodes.

During the eight days of which all the interaction at the display was coded, the display was in use 8.8% of its uptime and 1199 persons stopped to interact with it in 516 sessions. They were accompanied with (at least) another 202 persons, who only participated in viewing other people's interaction, without touching the display themselves. Figure 3 shows how the arrival of users was distributed across different days and times of day. As can be seen, use was slightly more active during the weekend and, in general, took place during in the evenings after working hours. Thus, most interaction was that of freetime users. However, the increased evening and night-time use can also be partially explained by lighting conditions more favourable to the display's visibility.

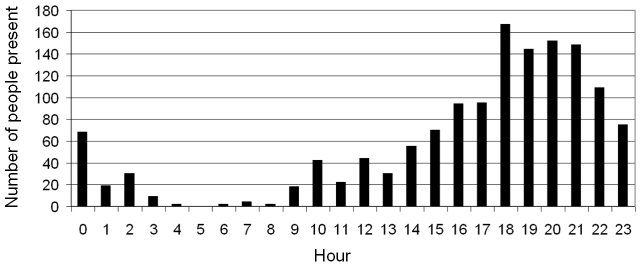

Figure 3. Number of people at CityWall at different hours. The values are sums from the eight days of use.

Only 18% of the users were individuals. The more detailed coding of the three selected days revealed more about the social configurations in multi-user situations. In multi-user situations, pairs were most common - they were present in 72% of these situations. Individuals and groups of three were seen more rarely in these situations (18% and 23%, respectively). Groups larger than three very rarely stopped at the display at any time.

Already, such a short analysis of statistics points to the social nature of a large display use in an urban environment. In the following sub-sections we analyze this theme in more detail, drawing from statistics and analyses of episodes of interaction.

Dynamics of Approach

The CityWall installation was set up along a busy public street. Logically, the first question is how people who pass by or go about their business there notice that there is an installation – or that the installation is interactive.

Noticing the Display

The presence of other users is important to the way new users arrive at the display. In 19% of the cases, CityWall was already in use by someone else when a new user entered the display and started using it. Given that the display was in use 8.8% of its total uptime, this indicates very sim-

Figure 4. Shelter from the rain

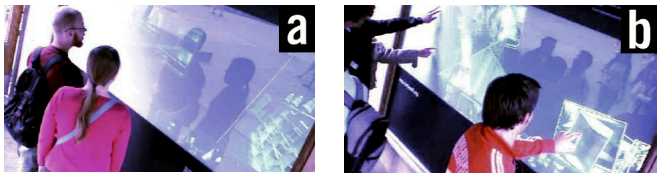

Figure 5. Stepwise approach to the wall

Figure 6. Parallel use

ply that seeing people using the display served as an attractor for more users.

In Figure 4, people are gathering under the sunshade to shelter from the rain that has just started. Despite the fact that objects on the screen are constantly moving, they are not paying attention to the screen, and stand with their back towards it just waiting for the rain to end. After ~20 seconds, a boy arrives there with his friends, notices the instructions next to the screen, and touches the screen, which, to his surprise, reacts. He utters "oooh", getting the attention of his friends and the older man standing next to him.

The example illustrates a typical pattern related to entries: *people most often notice the wall when someone is using it*. Visibility of the screen is not merely a sum of its physical properties. As the urban landscape is already full of visual clutter, people appear to be more attentive to other people's behaviour there. The user interviews supported this. Users commented that the system was hard to notice if nobody was using it and if one did not know what it was beforehand. They also stated that when they started using the installation it attracted a lot of attention from passers-by.

Stepwise Approach

After noticing the screen, people need to decide what to do about it. If there is nobody using it, or if there is room at the display, one can just step in and start exploring. From the reflection on the display, however, we observed that people often approached it in a stepwise manner [3].

In Figure 5a, the reflection shows the feet of two men observing it from several meters away. They wait for the couple in front to leave before approaching the screen. In Figure 5b, the reflection shows two women waiting for their turn just behind the person(s) using the screen. The latter layout is akin to queuing, but more about making one's presence visible to the one(s) using the screen than explicitly stating who is to take the floor next.

Interacting at the Display with Others

Multi-user interaction was the primary type of interaction observed at the display. In the most extreme case observed, there were as many as seven users touching the wall and browsing content with both hands at the same time! Because browsing content on the wall often had implications for the opportunities of others to use the wall, many turn-taking mechanisms were taking place.

Two baseline patterns of multi-user interaction were observed. Firstly, in *parallel use*, people can occupy an area of the screen and focus on their own task irrespective of the activities on their left or right. Alternatively, they can engage in *teamwork:* grouping with other users and focusing on the same object or set of objects. Failing to maintain the current organization, or to provide a smooth transition from one mode to another, leads to conflicts that then need to be handled separately.

Parallel Use

The CityWall screen is 2.5 meters wide, which means that it can accommodate several users at the same time. All users interviewed commented that the installation is most fun to use together and with one's friends. We observed that there are several ways people can organize parallel and joint activities at the screen.

The example above displays two instances of parallel uses. In Figure 6a, a group of more than ten young users has just gathered in front of the display and at the moment seven of them are trying to use it simultaneously using both their hands. Instead of coordinating their action, each of them tries to use the screen as an individual, not paying attention to what the others do. The end result is that nobody gets anything done as the screen receives too many inputs.

In Figure 6b, each of a pair of young males has picked his own photo, rotating and scaling it irrespective of what the other is doing. In this respect, their use also can be characterized as parallel. However, at some point, the two friends notice a similarity between their interactions, and start making fun of it. They start scaling up and scaling down the photos in a synchronized and pulsating fashion. This turns into a kind of dance, as the person on the left starts bending his knees and nodding his head according to the rhythm of the photos. This *dance between boys* lasts a brief moment only, as the man on the left cuts it short by taking a bunch of photos and throwing them on the one his friend is holding. The example shows how small the difference between doing things alone vs. doing things together can be.

Teamwork and playful activities

As in Figure 6b, it often happened that people who came to the screen with their friends did not just step in as individuals, but clearly teamed up in joint activities or started working on the same object.

In Figure 7a we see a large group of older tourists that have gathered in front of the display. Although the outset is similar to the one in Figure 6a, the social organization is different. In this case, not everyone goes to the display at once, but most of the group gathers behind the users, commenting and giving advice.

Figure 7. Teamwork

Teamwork is sometimes also a way of dealing with physical obstacles, or it can be adopted because it is more fun that way, or both. In Figure 7b, the two men are both holding a can of beer in their hand and, because of that (or inspired by it), they start scaling up the photo each grabbing one corner. Although CityWall was designed to enable two-hand usage, we observed many people using it with one hand only. As two-hand usage was not enforced (all moving, zooming and panning activities could also be done one handedly), this may have something to do with personal preferences, but not always. Not unusually was one-handedness due to a physical obstacle, as in the example above. It appears that people in downtown Helsinki are carrying all sorts of bags, skateboards, cameras, mobile phones or other items.

The photos displayed on CityWall were downloaded from public forums in the web. This resulted in the fact that an average user had no personal relationship with the content of the photos that happened to be on the screen when she appeared on site. Browsing one's own photos, should one have any online, or to more interesting photos was possible, but not well supported. This turned the user's attention from content to aspect of the interface. There were also users that seemed to take the content of the photos seriously, but a vast majority seemed to focus on playing with the interface. This was visible in the invention of *games* and different kinds of *nonsense activities* at the display. For example, people were playing *Pong*, throwing photos at each other, and *soccer*, building a goal out of two photos and trying to throw a third one in. There were also occasions when several people went to the screen bellowing and waving their hands irrationally.

Conflict management

Regardless of the type of activity individuals or teams are involved in, occasions where activities of different groups collide are likely to occur at some point.

In Figure 8a, the activities of the two groups conflict when the man on the left accidentally blows up a photo so that it goes on top of the photos that the group on the right was working on. They both turn their gaze towards the other group, and pull their hands out from the screen. In addition,

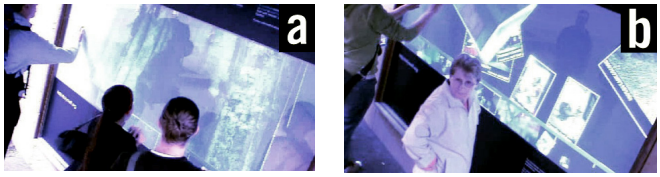

Figure 8. Conflict management by withdrawal

Figure 9. Social interaction inspired by conflicts

the woman places her hands in front of her chest making her withdrawal clearly visible to the other team.

In Figure 8b, an older woman has spent quite some time at the screen browsing the photos, and carefully scaling up and assembling some selected ones on the centre of the screen. At some point, two men start using the screen on the left, which soon leads into similar blow-up and overlap problem, preventing her from continuing. She turns to her husband (who has been watching the episode from behind) with a frustrated comment and bodily gesture, lifting her eyebrows and placing her arms on her hips. Instead of displaying her frustration to the other team, she seeks the support of the audience to make a moral statement about the situation. Similar observations have been made about responses to butting into a queue; the party feeling violated brings the attention of others to the observable problem, as if making the members speak in unison [13, p. 13–14].

Conflicts relate to the ownership of photos and their immediate surroundings, i.e. areas that may be needed for rotating, scaling and sorting the set of photos being worked on. The problem is that the UI causes people to unintentionally break these territorial borders, for example when photos are accidentally blown up or when using the timeline irrespective of what the other participants are doing. This was found the most disturbing conflict by the users interviewed. This is not to say that conflicts are always a problem – in the user interview, one user stated that friends helped out when something unexpected happened and it was actually fun when photos got accidentally and unexpectedly too big.

Although conflicts take place, they can also have positive consequences for the social organization at the display. In Figure 9a, the boy starts to take over the photo on the top left corner of the screen. The man in white shirt steps in claiming: "It is mine, don't touch". The participants take this as a form of joking and laugh together. In Figure 9b, the man on the left has accidentally thrown a photo on top of the one the couple on the right is working on. After a joint recognition of the conflict, he and his friend start throwing more photos at the other group. The man on the left responses "bravo bravo", all four laugh together, and the group on the left withdraws, handing over the floor to the couple.

Transitions between activities and participants

Above we have shown that CityWall supports various joint activities, such as *browsing* and *scaling* of photos, playing *Pong* or *soccer*, or even *dancing with photos*. Similarly, we have shown how people user the screen together with others

in various combinations, and how they negotiate who gets to, or should use, the screen, and when.

The possible activities or possible combinations of people are as infinite as the imaginable contents of photos displayed on the screen. Rather than telling what all the possible activities supported by CityWall are, we show how the management of transitions of different kinds is intertwined with the physical interaction with the display and with the other users, as well as sense-making of photos.

Floor and Turn-Taking

Research on ordinary conversation has shown how the participants monitor the current speaker and orient to transition relevant places (TRPs), i.e. moments when it is possible to take the floor [20]. Also the speaker recognizes these windows of opportunity, and has the means to select the next speaker slot or ways to continue keeping the floor across a TRP [20]. Similarly, the users, when giving the floor to others, could provide for fluent transition by making their withdrawal noticeable. For example, they could leave the screen throwing a photo or fast-forwarding the timeline. Another example of *terminal activity* was when people, just before their exit, slowly but steadily moved towards the side of the screen haphazardly poking at the elements of the UI. The conclusive nature of the activity is visible in how they play with whatever happens to be visible on the screen, with no attempts to bring in any new items [21]

By observing the actions of others, people can anticipate when it is appropriate to go and take the floor [20]. On one occasion, a boy who came to the installation with his mother made a move towards the screen when there was only one person using it and thus plenty of room. His mother however prevented this, ordering "*noo-no no, wait it's their turn now.*" This shows that appropriate moment for entry or transition between users is not a matter of available space at the display, but a result of a more complex reasoning and negotiation between the participants.

Expressive and Pondering Gestures

Should one want to keep the floor, one should take into account that others may possibly jump into any *idle moment* or *transition relevant place*.. Also, should one want to engage the other party in interaction, one might have to wait for a suitable moment to do so.

We observed that people can use *distinct ways to touch or hold photos* in order to serve the management of transitions between users or activities. Furthermore, transition management of this kind is not a separate activity, but intertwined with cognitive and physical aspects of use.

In Figure 10a, the woman on the left is carefully moving objects around the left side of the wall. In contrast, the couple on his right is performing scaling up with grandiose gestures, on the verge of entering her personal space. In Figure 10b, the man on the left is holding a photo in his hand, keeping it in constant small movement, waiting for

Figure 10. Pondering grip vs. grandiose gestures

the right moment to interfere with his friend's intense interaction with the photo on the right. When the right moment comes, he proposes "catch", after which the two start throwing his photo back and forth (cf. Pong playing above).

When it comes to holding and manipulating photos, the *intensity of touch* can vary a lot. *Grandiose gestures* provide for an intensively tangible interaction experience that also communicates to the other participants. On the other hand, we recorded many events where people were holding photos with a *pondering grip*, as if thinking of what to do with the photo or waiting for an inspiration or action of a co-participant that would open an opportunity of some kind.

Concluding Actions

When people browse and play with photos together, they use verbal and physical means to communicate and ensure that they have a shared point of attention (e.g. a photo or set of photos), as well as a common understanding of the frame of activity, i.e. what to do with the object.

Before changing to a new object or frame of activity, it is natural to summarize the earlier ones, by, for example, saying that something was *fun* or *cool* or *boring* etc. Assessments can also look forward in time. For example, one can establish a new point of attention pointing at an object saying "*ooh*" or "*hey look!*"

Although possible in theory, it would be difficult, i.e. socially obtuse, to go to the screen and use it with someone without presenting opinions or assessments at some point. For example, when leaving the site, people sometimes leave their fingerprint or make a mark of some kind.

In Figure 11a, just before leaving the screen, the man scales down and arranges all visible photos in a gallery-like layout, wasting no space. In Figure 11b, the man entertains his audience by blowing up one photo to fill the whole screen and announcing in a loud voice "the world is MINE!"

There are several ways of leaving a mark. At exit, people can, for instance, give momentum to the timeline or desktop so that photos fly there for a moment, or they can select a funny or embarrassing photo to leave on top.

Figure 11. Leaving a mark

Roles and Social Configurations

As said earlier, 18% of the use episodes contained only one person interacting with the display. Of the complementary 82%, 20% of the time there was more than one group present at the wall. Thus, in total 0.82 * 020 = 16% of the use situations at the display took place when the display was interacted with by people who were strangers. The introductory chapter already stated that the most common group sizes were two, three and one (in this order).

When people team up at the screen, in principle they have equal rights to interact with it. However, in our study, instead of uniform orientation, individuals in groups often oriented to complementary roles or social configurations. The most frequently recurring social configuration was the *teacher–apprentice* setting, where one or more users took the role of an experienced or technologically savvy user, and went on to explain the features of the application, assisting the other members of the group when needed.

In Figure 12a, the man shows and tells his girlfriend how the touchscreen works. The posture of his girlfriend clearly tells she does not intend to touch the screen yet. In Figure 12b, the man in front of the display gives an overview of various parts of the UI. For that purpose, he has positioned himself between the screen and audience. In the user interviews, it also came up that it was easy to learn to use the touchscreen just by following the example of others, and nobody actually needed to read the instructions printed next to the installation.

Another role is that of a *comedian*. For example, the man appearing in Figures 9a and 11b looked actively for opportunities to entertain his audience. Role taking is also an essential part of gaming; when playing *pong*, we are tied to the fuzzy set of rules that relate to the game, and do not, for example, interact with the timeline. Although we did not find instances of authorship in the sample, it is easy to imagine roles that relate to the content of the photos; for example, the photographer or subject of a photo is likely to highlight an aspect of it different from the one an average passer-by might highlight.

Role-taking can also be seen as a way to deal with the complexity of the social setting and usability problems it causes. When several people gather at the display, it is not feasible to assume that all could step in as the main operator of the system. As there was no concrete support for queuing or turn-taking, people often filled in any space that opened in front of the screen. Different types of social configurations at the display make *it possible for multiple participants to act* at once. For example, when a person is interacting with the wall, her friend can adopt the role of an assistant or a commentator, affecting the course of events without having to touch the display. Casting was not only done at the outset, but people changed places during their interaction. Supported by verbal reports, people were also able to align their parallel and joint activities [12].

Certain rights and constraints apply also to social configurations between strangers. Unacquainted persons *need a reason* to enter face encounters with each other in public places [6, p 124]. In the case of CityWall, conflicts between parallel tasks of two or more users or teams were the main reason for interactions between strangers. Users did try to avoid interfering with parallel activities, but the system did not support *the norm of social segregation* between the unacquainted, but made photos accidentally inflate or fly across the screen. This then forced the users to engage in conflict management with each other. The positive outcome is that the system can make strangers interact with each other. However, we should also think of other means to support this, not rely on positive effects of accidental and unwanted system features.

DISCUSSION

Encounters and collective interactions at the display

Unlike in most of the settings in which public displays have been studied in previous research, a real urban environment is populated by individuals and groups that are strangers to each other. A striking result was how these people were configured in groups of users and crowds of spectators rather than as individual users. They were able to use the display both in parallel and collectively by adopting different roles. That the use of the display was highly non-individualistic was evident both in statistics and in the analysis of the selected episodes of interaction. Pairs stopped at the display more often than individuals and, as a result, only 18% of use sessions consisted of only one user.

We believe that *learning from other users* may be one of the key explanations for this. First of all, seeing someone using the display made people aware that it was not just another shop window, but an interactive installation. Also, when moving closer to the display, when standing behind the earlier users, and when actually using the display, people learned more about its interactive properties.

Both stepwise entries—queuing for one's turn—and more direct entries—occupying a space while trying not to disturb the existing users—were observed. Previous research (e.g. [3]) has reported similar cases of increasing numbers of people congregating around a display, but has not described in detail the patterns of engagement taking place between the people.

Interaction as a Performance

Content on the wall and features of the interface were used as resources to coordinate the activity and to create events

Figure 12. Teacher-apprentice setting

or interactions so they were *meaningful in front of others*. For example, interactions like photo-moving and scaling turned into games like *Pong playing*. Some gestures were *made* salient to others. "*Grandiose gestures"* and "*pondering grips"* were used to manifest the volume of the user's actions and her intentions towards others, while also marking the boundaries of the workspace that the user felt she had claimed as her own. Also the presence of strangers—all the other people walking past the installation, sometimes stopping by to observe what the users are doing—has an effect on one's activities at the CityWall, which can be perceived as a performance in the city space.

The contribution of this study to "interaction as performance" [11] is to evidence how the size of the CityWall created a sufficient space for a "stage" for multiple users. As shown, users were able to adopt different roles, such as being teachers, apprentices, clowns, or members of the audience. In some cases, multiple activities were taking place at the same time at the display. Such *a*symmetric participation patterns have not been reported to this extent before. This shows yet another aspect of how a public display configures spaces and surfaces [17] in support of social interaction. The multi-touch feature of the interface was central, as it supported expressive gestures that helped participants in coordinating, communicating and acting out different roles.

Restructuring the Space

In this study, we installed a multi-touch and multi-user display in a public urban space. While experimentation of "placing" public displays has been considered in previous work [1, 3, 8], it has focused mostly on private and semi-public settings. Previous work has also been free of theoretical considerations, such as how these interventions in urban settings become also architectural ones, restructuring the space and changing the way people behave around these installations.

CityWall was located in a central pedestrian area between the bus and railway stations and important shopping centres. The space connects key locations in the city and contains several small shops and cafes. The space, like similar ones in other cities, is transformed during events. It hosts temporary attractions and gathers groups of passers-by.

In our observations and interviews, we could see how the space is used by the widest variety of people that are strangers to each other. To repurpose an existing architectural element in this area we chose to install the CityWall on an existing shop front window, which was turned into an interactive display. This solution was very different from introducing in the space a new architectural element as a construct of its own. For example, "a box" would not only be more visible but would also change how people could move in the space.

Our architectural solution invites passers-by not so much because of its physical properties, but because it taps different cultural references. As argued by architect Bernard Tschumi, we should not look at architecture as an object, but as "interaction of space and events" [23, p. 162]. People have been grouping around shop windows in situations before, for example, to shelter from the rain (Figure 4), but new technology can be used to extend the bodily presence of people, and change how "bodies … generate spaces produced by and through the movements" [23, p. 154]. It was not our intervention that created a new place but the people themselves with their appropriation of the space surrounding the display.

CONCLUSIONS

To conclude, we have shown how a large multi-touch screen can create a stage and therefore a place for strangers to come into contact; for example, users at the display may attract other users. However, strangers acted mostly separately, but courteously, in parallel, and interacted with one another mostly after a conflict. As previously observed in studies on the use of tangible interfaces, users have the opportunity to engage in performative interactions [11]. The particular size of the screen used in this research and its multi-touch feature supports bodily interactions with the display, making it possible to be expressive towards other participants, and helping them to take up roles and negotiate turn-taking as well as different kinds of collaborative activities. Interactions between an individual and the system (the core of CHI) were turned into expressive acts toward the other participants.

Finally, by analyzing interaction in a detailed way, it becomes clear how the public availability of such an interface and of its digital objects created a tension between personal space and action and publicly available resources. We think we are still in the infancy, not only of our own project, but also of the whole application area of designing applications for such group usages. In particular, we are thinking of established "norms" of conduct that apply to other "older" publicly available objects. We find it important to think about design separately for (small or large) groups of users versus individual users. Design should support performative acts and facilitate asymmetric and ad hoc role-taking, thus letting users learn the opportunities for interaction from their peers. Previous work has highlighted the importance of the display to communicating what it offers an individual user directly [3], but our study suggests that there may be ways to support *social learning* in relation to what it offers. For one, CityWall's large physical size appeared to support making interactions visible to others both gesturally and as effects on the display when this was wanted. Large display size and visibility also supports immediate availability of content to interact with.

ACKNOWLEDGMENTS

This work has been funded by the 6th Framework Research Programme of the EU, through the IPCity project (FP-2004-IST-4-27571). We would also like to thank Vikram Batra, Lasipalatsin Mediakeskus and City of Helsinki Cultural Office for their help.

REFERENCES

1. Agostini, A., De Michelis, G., Divitini, M., Grasso, M.A. and Snowdon, D. Design and deployment of community systems: reflections on the Campiello experience. *Interacting with Computers 14,* 6 (2002), 689–712.
2. Ananny, M. Strohecker, C. and Biddick, K.. Shifting scales on common ground: developing personal expressions and public opinions. *International Journal of Continuing Engineering Education and Life Long Learning 14,* 6 (2004), 484–505.
3. Brignull, H. and Rogers, Y. Enticing people to interact with large public displays in public spaces. In *Proc. INTERACT 2003*, IOS Press (2003), Amsterdam, 17–24.
4. Brignull, H., Izadi, S., Fitzpatrick, G., Rogers, Y., and Rodden, T. The introduction of a shared interactive surface into a communal space. In *Proc. CSCW 2004*, ACM Press (2004), New York, 49–58.
5. Fatah gen. Schieck, A. and Kostakos, V. Exploring digital encounters in the city. Position paper presented in CHI2007 workshop on Shared Encounters. Available at www.cs.bath.ac.uk/~vk/files/chi07.pdf
6. Goffman, E. *Behavior in Public Places*. Free Press (1963), New York.
7. Huang, E.M., Mynatt, E.D. Semi-public displays for small, co-located groups. In *Proc. CHI 2003*, ACM Press (2003), New York, 49–56.
8. Huang, E.M., Mynatt, E.D. and Trimble, J.P. Displays in the wild: understanding the dynamics and evolution of a display ecology. In *Proc. Pervasive 2006*, Lecture Notes in Computer Science 3968, Springer (2006), Berlin, 321–336.
9. Huang, E.M., Mynatt, E.D. and Trimble, J.P. When design just isn't enough: the unanticipated challenges of the real world for large collaborative displays. *Personal and Ubiquitous Computing* 11, 7 (2007), 537–547.
10. Izadi, S., Brignull, H., Rodden, T. and Rogers, Y. Underwood, M. Dynamo: A public interactive surface supporting the cooperative sharing and exchange of media. In *Proc. UIST 2003*, ACM Press (2003), New York, 159–168.
11. Jacucci, G. *Interaction as Performance. Cases of Configuring Physical Interfaces in Mixed Media.* Oulu University Press (2004), Oulu, Finland.
12. Kurvinen, E., Lähteenmäki, M., Salovaara, A. and Lopez F. Are You Alive? Sensor data as a resource for social interaction. *Journal of Knowledge, Technology & Policy* 20, 1 (2007), 39–49.
13. Livingston, E.C. *Making Sense of Ethnomethodology.* Routledge & Kegan Paul, London, 1987.
14. Matsushita, N. and Rekimoto, J. HoloWall: designing a finger, hand, body, and object sensitive wall. In *Proc. UIST 2003*, ACM Press (2003), New York, 159–168.
15. Morris, M.R., Huang, A., Paepcke, A. and Winograd, T. Cooperative gestures: multi-user gestural interactions for co-located groupware. In *Proc. CHI 2006*, ACM Press (2006), New York, 1201–1210.
16. Robertson, T. The Public Availability of actions and artefacts. *Computer Supported Cooperative Work 11,* 3–4 (2002), 299–316.
17. Rogers, Y. and Rodden, T. Configuring spaces and surfaces to support collaborative interactions. In O'Hara, K., Perry, M., Churchill, E. and Russell, D. (eds.). *Public and Situated Displays: Social and Interactional Aspects of Shared Display Technologies*. Kluwer, Dordrecht, The Netherlands, 2003, 45–79.
18. Rogers, Y., Hazlewood, W., Blevis, E. and Lim, Y.-K. Finger talk: collaborative decision-making using talk and fingertip interaction around a tabletop display. In *Proc. CHI 2004*, ACM Press (2004), New York, 1271–1274.
19. Russell, D.M., Drews, D. and Sue, A. Social Aspects of Using Large Public Interactive Displays for Collaboration. In *Proc. UbiComp 2002*, Lecture Notes in Computer Science 2498, Springer (2002), Berlin, 229–236.
20. Sacks, H., Schegloff, E.A. and Jefferson, G. A simplest systematics for the organization of turn taking in conversation. *Language* 50, 4 (1974), 696–735.
21. Schegloff, E.A. and Sacks, H. Opening Up Closings. *Semiotica VIII*, 4 (1973), 289–327.
22. Skog, T., Ljungblad, S. and Lars Holmquist, L.E. Between aesthetics and utility: designing ambient information visualizations. In *Proc. InfoVis 2003*, IEEE Computer Society (2003), New York, 233–240.
23. Tschumi, B. Architechture and Limits III (1981). In Nesbitt, K. (ed.), *Theorizing a New Agenda for Architecture. An Anthology of Architectural Theory 1965-1995*. Princeton Architectural Press (1996), New York, 156–162.
24. Vogel, D. and Balakrishnan, R. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proc. UIST 2004*, ACM Press (2004), New York, 137–146.
25. Wisneski, C., Ishii, H., Dahley, A., Gorbet, M., Brave, S., Ullmer, B., Yarin, P. Ambient displays: turning architectural space into an interface between people and digital information. In *Proc. CoBuild 1998*, Lecture Notes in Computer Science 1370, Springer (1998), Berlin, 22–33.
26. Wu, M. and Balakrishnan, R. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proc. UIST 2003*, ACM Press (2003), New York, 193–202.