Gesture Elicitation Study on How to Opt-in & Opt-out from Interactions with Public Displays

Isabel Benavente Rodriguez, Nicolai Marquardt
University College London
{isabel.rodriguez.15, n.marquardt} @ucl.ac.uk

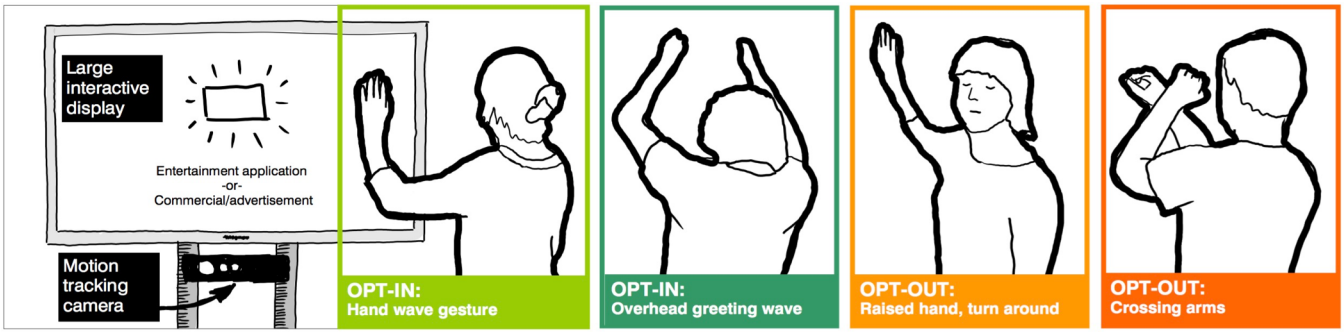

Figure 1. Overview of user defined gestures for opting-in and opting-out from interactions with public displays.

ABSTRACT
Public interactive displays with gesture-recognizing cameras enable new forms of interactions. However, often such systems do not yet allow passers-by a choice to engage voluntarily or disengage from an interaction. To address this issue, this paper explores how people could use different kinds of gestures or voice commands to explicitly *opt-in* or *opt-out* of interactions with public installations. We report the results of a gesture elicitation study with 16 participants, generating gestures within five gesture-types for both a commercial and entertainment scenario. We present a categorization and themes of the 430 proposed gestures, and agreement scores showing higher consensus for torso gestures and for opting-out with face/head. Furthermore, patterns indicate that participants often chose non-verbal representations of opposing pairs such as 'close and open' when proposing gestures. Quantitative results showed overall preference for hand and arm gestures, and generally a higher acceptance for gestural interaction in the entertainment setting.

Author Keywords
Public displays; gesture interfaces; elicitation study; user-defined gestures; whole-body interaction

ACM Classification Keywords
H.5.m. Information interfaces and presentation

Please cite as:

Isabel Benavente Rodriguez and Nicolai Marquardt. 2017. Gesture Elicitation Study on How to Opt-in & Opt-out from Interactions with Public Displays. In Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS '17). ACM, New York, NY, USA, 32-41. DOI: https://doi.org/10.1145/3132272.3134118

INTRODUCTION
Combining public interactive displays with body- and gesture-tracking cameras opens up new possibilities for engaging interaction experiences. Examples include: public ambient displays operated through hand and body gestures [38], interactive whiteboards reacting to people's proximity [18], responsive interactive floors [36], and even proxemic-aware advertisements [39]. Often these systems include depth-sensing cameras (such as the Microsoft Kinect [19]) to recognize people in front of the large display, and to track their motion and gestures they perform (which is then interpreted as input for the system). Previous studies explored the design space for gestural interactions with such large displays [26, 32, 34, 37, 38] and refined the vocabulary for new gestural interactions. However, one challenge with public interactive installations in particular is that passers-by often do not have a *choice* to engage voluntarily [5], as the motion-tracking cameras of these systems are running continuously. This means that in most cases any person walking by a large interactive display or on an interactive floor is directly interacting with the system, which could lead to unintended interactions that are confusing, frustrating, or even embarrassing for a person walking by [4, 6, 30]. This also relates to the notion of *dark patterns* [14], where Greenberg et al. discuss interactive systems that (possibly unintentional) "*violate social mores*", resulting in a bad user experience. This led to concrete suggestions [5] addressing this problem, such as clearly communicating interaction zones or – as we do in this paper – "*making opt-out easy […] for example, through explicit gestures*" [5].

With our research study, we explore how people could use gestures and voice commands to explicitly *opt-in or opt-out of interactions with public installations*. We are interested in identifying gestures for initiating and stopping interactions with a system, and therefore giving people a choice

when interacting (or not interacting) with public interactive systems. We conducted a *gesture elicitation study* to investigate how people can use gestures to opt-in or opt-out from interactions with an interactive system, such as large public displays (Figure 1). The aim is to explore ways that allow people to be more explicit when expressing their consent or refusal to interact with a body-tracking system.

For our study, we invited 16 participants (in 8 pairs of two) to propose gesture commands representing 'opting-in' and 'opting-out' using successively four different parts of their body (either only head, hand, arms, body) or by using their voice. Because we focus on interactive systems that react to people's presence/gestures from a distance (e.g., [29, 36, 38]), we do not include direct touch-screen interactions (which, in many of the proposed scenarios, would not be possible to perform due to different reasons: screen out of reach, behind a window, screen is an interactive floor, etc.). In combination with two scenarios we presented to the participants (commercial vs. entertainment setting), this resulted in a list of 20 referents (2 x 5 x 2). The design of our study was informed by the insights of previous gesture elicitation studies in HCI [9, 26, 34, 35], including the calculating agreement scores for proposed gestures [40] and the revised calculations (and toolkit) [33].

With our paper, we contribute the categorization of the 430 proposed gestures, including the calculated agreement scores indicating higher consensus for torso gestures and for opting-out with face/head. We also discuss in detail five interaction themes across the proposed gesture sets that we observed during the experiment. Patterns indicate that participants often chose non-verbal representations of opposing pairs such as 'close and open' when proposing gestures. Quantitative results showed overall preference for hand and arm gestures, and generally a higher acceptance for gestural interaction in the entertainment setting. We discuss how the suggested gestures and our observed patterns can inform a future interaction vocabulary for opting-in or opting-out of interactions with gesture-based systems. Finally, we also address topics related to interactions in public environments such as comfort, appropriateness, and privacy concerns.

RELATED WORK

Interactive technologies are increasingly available in public environments: from art installations visualizing real-time data [16], interactive floors [22, 36], installations allowing viewing delicate objects for example in museums [30], systems to buy goods [12], to an Information Wall [23]. Many of these systems include some kind of tracking technology (e.g., camera based) to recognize people and any gestures they perform with their hands, arms, or full body as input for the system to allow interaction.

Along with the increasing availability of technology installations in public spaces, research has made efforts to understand not only our interactions but the impact these interfaces have on people's behavior. The insights contribute to the improvement of their design by addressing the visibility of interfaces [10, 27], favorable environmental and spatial factors [11, 17], and the likelihood of engaging potential users or ways to hold [1, 6] or regain [39] attention.

According to Preece et al. [31], one of the main challenges when designing air-based gestural input is "*to consider how a computer system recognizes and delineates the user's gestures*". Similarly, Bellotti et al. [4] identified 'Addressing' a system as one of the five key questions for designing ubicomp interactions: "How do I address one (or more) of many possible devices?". The problem here is to define the beginning and end of a movement and teaching the system to notice the difference between a natural 'side effect' movement and the specific intentional movements aiming to operate the system. For the particular application with public displays, Walter et al. [38] developed StrikeAPose to investigate how to best reveal mid-air gestures for interactive public displays to users. Their focus was on initial gestures that served as gesture registration — the starting point of an interaction. Previous studies investigated the design of gestural interaction for many different contexts, such as body motion and gestures to interact with public displays [37], foot gestures as commands [2], or collaborative gameplay on public displays [29].

To find appropriate gestures that fit the different interaction contexts with technology, research has been done using *gesture elicitation studies* that put potential users at the center of gesture design and leverages their proposed gestures to inform interaction design. Gesture elicitation studies show participants the desired result of a command (called *referent*) and asks them to suggest a gesture (called *symbol*) that would trigger it [25]. In recent years, elicitation studies have been applied to different application areas. For example, it was used to categorize multi-touch gestures for surface computing [41], for finding methods for interacting with interactive TV [34] and identify single hand micro gestures [7]. In multi-device interaction, gesture elicitation studies are used to find ways to pair and interconnect devices [8] and to identify gestures in multi-display environments for transferring information between devices [32]. For large public displays in particular, recent studies explored how gesture control can be used for interactive games [9] or for using internet browser functions on a large display through gestures [26].

These gesture elicitation studies have in common that they helped to identify patterns among their participants' gesture proposals that can inform the development of future systems. Combined with metrics such as acceptance of a gesture, recall likeliness, or agreement among participants it is also possible to find out whether certain proposed gestures likely fit to the expectations of a broader user group and hence user-friendly interactions. Wobbrock et al.'s [40] formula helps to calculate said agreement scores, however, some previous studies found that agreement scores for the proposed gestures remained relatively low [32]. Vatavu and Wobbrock [33] later provided a toolkit advancing the sys-

tematization of elicitation analysis by introducing measures that capture disagreement rates and co-agreement rates, as well as a statistical significance test that allows to compare agreement rates.

Informed by those previous elicitation studies [9, 26, 34] and matching analysis techniques [33, 40], we apply this method to gestural interaction with public displays. In particular, we explore what kind of gestures a person can perform to start, refuse or abandon an interaction.

OPTIN+OPTOUT GESTURE ELICITATION STUDY

In this section, we describe our gesture elicitation study design for opting-in and opting-out of interactions with public displays.

Participants

A total of sixteen participants (8 female, 8 male) took part in the elicitation study with ages ranging from 22 to 37 (M=27.4, SD=4.1). Participants were students with varied academic backgrounds (n=5) or professionals, such as industrial designer, developer (n=3), engineer, or IT consultant (n=2). They undertook the study in pairs and were either partners or friends. The participants' technology proficiency was relatively high: fifteen of the participants use a smartphone and laptop daily and eight use at least three devices (smartphone, tablet, laptop) every day. With only one exception all participants have tried a motion-capturing video game before, such as games for Nintendo Wii or Microsoft Kinect. Moreover, 44% have used a gesture installation in public before.

Study Design

The setup and design of our study is directly informed by similar gesture elicitation study designs in HCI (e.g., [34] and [26]). To avoid a quiz-like atmosphere with participants feeling put on the spot, the subjects were invited to take part in the study in pairs with someone familiar. As discussed by Morris et al. [26], working with participants that know each other creates a relaxed environment where the tasks given become stimuli to brainstorm as a team rather than a challenge demanding instant answers.

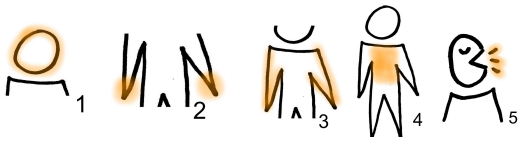

Figure 2. Overview of body part constraints: 1. face & head, 2. fingers & hands, 3. arms, 4. torso/posture, 5. voice.

The groups of two were presented with two scenarios, each introduced by a short video clip of a public installation, and were then asked to come up with gestures expressing their consent of opting-in to an interaction or opting-out. We minimized the influence of the videos on the user defined gestures by making sure the videos only showed mid-use scenes, not the moment of approaching and noticing the installation. The first scenario we presented to participants is a commercial setting, and the installation of the second scenario was for entertainment and recreational purposes. After a freestyle category, constraints were given allowing the participants to only use certain body parts for proposing gestures (see Figure 2). We selected those body-part constraints to closely match the capabilities of currently used technical tracking systems used in public installations (e.g., cameras that only track people's faces, hands, arms, their full body (e.g., depth-sensing cameras), or only voice).

For each body part category, the participants were encouraged to consult with each other and come up with as many gestures as they wanted (but with at least two). To move on to the next category, the pair had to agree on a favorite and declare it consensus of the group. 'Agree to disagree' was also a valid submission.

The sessions were video-recorded for later thematic analysis. A post-study questionnaire using Likert scales asked the participants to rate their level of agreement for a set of statements evolving around topics such as public comfort or privacy concerns for both commercial and entertainment scenarios. The questionnaire also included open text questions about suggestions on benefits and limitations of gesture-based installations in public spaces.

Materials, Referents and Procedure

The sessions took place in a quiet study room at a university campus. The furnishing was laid out to offer both participants enough space to comfortably stand in the room and move around. Sessions started with the participants filling out a pre-questionnaire collecting demographic data and experience levels. During the session, the participants were shown edited 30-second versions of the videos with highlight sequences to demonstrate both scenarios (videos are YouTube videos from [20, 21]). A 55-inch TV screen showed both the introduction videos as well as pictures of a mirror (first scenario) and an illuminated cube (second scenario) with occasional animations to give the participants a reference point to address their gestures to and support their imagination.

Commercial scenario: For the commercial scenario, an interactive mirror in a clothing store was presented that gave its users a preview of how different garments would fit them. The user performed body gestures to browse through a catalogue and select items from the given categories. The mirror also reacted to voice commands. Overall, the first scenario represented a commercial context were gesture-based installations can be used to carry out a certain marketing agenda, influence purchase decisions, or place personalized advertisement.

Entertainment scenario: The second scenario was also introduced by a video that showed an interactive cube installation (approx. 3x3 meters) which imitated the dance rhythm of a user through colorful animations on its surface. The cube represents the sort of playful public installation whose purpose is to entertain and engage its users.

		Referent	Proposed gesture
COMERCIAL	Opt-in	Face&head	Nodding, smile, long blink, **move head from side to side,** wink, draw a circle with the head, bow head, stick tongue out, shake head, raise eyebrows, scan own body with eyes, make long eye contact, nod up (once), kiss, move head back and forth (= cultural gesture: like a dance move), turn head to one side
		Fingers&hands	Wave, thumbs up, **opening fist,** clap, press a start button with hand as cursor, peace sign, close fist, slide with one finger, ‚twinkle' fingers, punch forwards, shoot with index finger and thumb, snap fingers, put fingers in heart shape, palms in front of body and turning down
		Arms	**Swipe, bigger greeting wave, circular arm movement in front of body,** pull apart parallel palms, wave with crossed arms over head, shrug, pressing button with arm as cursor, dance with arms (cultural gesture: ‚Chicken Dance')
		Torso&posture	Firm stand in front of it, **twist upper body, sway from side to side,** align with cut out on screen, spin around (once), bow, slouch, wiggle body, sway back and forth, stand in designated area (marked on the floor)
		Voice	„Hello/Hi", **„Hello mirror/name of mirror", „Show me clothes",** „Mirror, mirror on the wall", „Turn on", „Activate", „Let's look at t-shirts", „Wake up", „I want to try clothes on", „What do you have for me?", „Open", „Yo", „Start", „What's up?", click one's tongue
	Opt-out	Face&head	Shaking head (No), long blink, turning head round and round, nod, bow, chin up, blow
		Fingers&hands	**Closing fist,** press button wit hands as cursor, show palm (cultural gesture: Stop), hand flick, gesture: shoo flies away, erasing/wiping, clap, flick away with index finger and thumb, swaying index finger (cultural gesture: No), cross ('X') both index fingers in front of body, thumbs up, thumbs down, snap fingers
		Arms	Pulling lifted arm(s) down, make a cross ('X') with both arms in front of body, gesture: shoo flies away, erasing/wiping, circular movement with both arms (cultural gesture: like director of an orchestra), push vertically lifted lower arms together, press button using arm as cursor, knock, cross arms (cultural gesture: Cut), gesture: like arming a ball away
		Torso&posture	Turning away and showing back for a few seconds, **crouch/slouch forward, bow,** turn half away (45°), squat, shrug, shake "it off"
		Voice	"Goodbye/Bye", **"Stop",** Close", "Done", "Finish", "End", "Exit", "Safe my information", "I'll be back", "Thank you", "See you", "Shhh", whistle
ENTERTAINMENT	Opt-in	Face&head	**Move head from side to side, move head around,** nodding, smile, headbang, blow, make eye contact, headbutt, "Pick-a-boo" (hiding face behind hands), stretching grimace, kiss, wink, raise eyebrows, shake head, move head back and forth (= cultural gesture: ‚The Egyptian' dance move)
		Fingers&hands	**Magic fingers/spirit fingers, Jazz Hands,** Clap, small wave, point at installation, shadow puppet, snap fingers, blow kiss with hand, greeting gesture: index and middle finger on forehead (symbolic lifting sb.'s head), "blinking" fingers, draw circle with finger, swipe, opening fist, gesture: director of an orchestra, 'The Queen' wave, waving with two hands
		Arms	**Big arm wave, big circular movement of both arms, waving with two arms over head (like rescue),** arm wiggle, lower arms lying horizontally on top of each other, paddling with both arms (like dog), arm wave from one arm to the other (dance move)
		Torso&posture	**Twist hips, twist upper body, body wave/sway,** firm stand, any dance, come close and touch surface, spin, side to side stretching of upper body (like warming up pre-exercise)
		Voice	"Hey/Hello", **start singing, whistle,** "Start", "Interact with me", "What can yo do?", "Wakey, wakey", "Knock, knock", "On", "Bam", "Shazam", click one's tongue, "Go", "Ooohh" (sound), "Let's dance", "Let's rock 'n' roll", "Let's do it" "What's up", blow
	Opt-out	Face&head	Shaking head (No), **head down for a few seconds,** scrunch face, blow kiss, nop up (once), long blink, wink, look up with entire head
		Fingers&hands	**closing fist, show palm (cultural gesture: Stop), small wave,** hand flick, 'shut' hands in front of face, cross hands (cultural gesture: Cut), index finger on lips, snap with fingers, wipe something from shoulder, flip hands to back of the hand side, hinted bowing gesture with hands
		Arms	Wave with both arms, cross ('X') lower arms in front of body, cross arms (cultural gesture: Cut), fold arms, pulling lifted arms down, one clap, push parallel palms together, push palms forward, "gather" animated cubes from the installation and push them down
		Torso&posture	Turning away and showing back for a few seconds, **turn half away (45°),** bow, crouch, spin, hunch one's shoulders, standing still, wiggle/twist hips
		Voice	"Goodbye/Bye", **"Thank you",** "Finish", "See you", "Done", "Stop", "Shut down", "Laterrr", "I'm leaving", "End", "Shhh", hiss, blow

Table 1. Summary of the proposed gestures from the gesture elicitation study (most frequent gestures mentioned at least three times are highlighted in bold, while gestures in green were proposed at least five times).

Following the videos, the experimenter explained how the installations work to capture people's interactions with cameras. The participants then started proposing gestures for 'opting-in' and 'opting-out' in this scenario, by both verbally describing and demonstrating the gestures in front of the display.

For each constrained category as well as the freestyle category, the pairs had to agree on their favorite gesture which was then declared the group's consensus. It was also allowed to nominate a 'best gesture' that was previously mentioned in the freestyle category provided it matched the body part requirement. This element was implemented following the example of Morris et al. [26] in order to obtain very basic indications to weight the proposed gestures against each other.

For the actual elicitation, a set of 20 referents were used that comprised of the multiplication of body part challenge (5), opt-in/out situation (2), and scenario (2). The sessions concluded with a post-study questionnaire where participants rated their level of comfort when performing gestures with the different body parts that were explored through the given limitations (using 5-point Likert scales). Last, the participants were asked about their opinions and perception regarding gestural interactions in such scenarios.

FINDINGS

The results of the elicitation study can be divided into quantitative and qualitative results. The former begins with an overview (out of the full set of 430 proposed gestures) and shows how these are distributed over the body part categories as well as the different scenarios. Then agreement scores highlight which of these categories prompted the highest scores. We also review how the favorites nominated by each group of participants relate to their overall mentioned frequency. In the qualitative part of the elicitation findings we will discuss patterns that emerged among the gesture proposals. The chapter concludes with the evaluation of the questionnaires.

Complete Set of Participants' Proposed Gestures

The eight groups of participants proposed 430 gestures in total. Table 1 shows a summary (out of the full set of the 430 distinct gestures) that were proposed by the participants. The most frequent gestures that were mentioned at least three times are highlighted in bold, while gestures in green were proposed at least five times. In order to create an expressive overview, we grouped similar gestures into

the same category: for example, 'one-arm waving with wrist movement' and 'one-arm waving with elbow movement' were eventually summarized to 'waving'. In the set of proposed gestures, several similar gestures were proposed for different functions (e.g. the same gesture such as waving was proposed multiple times).

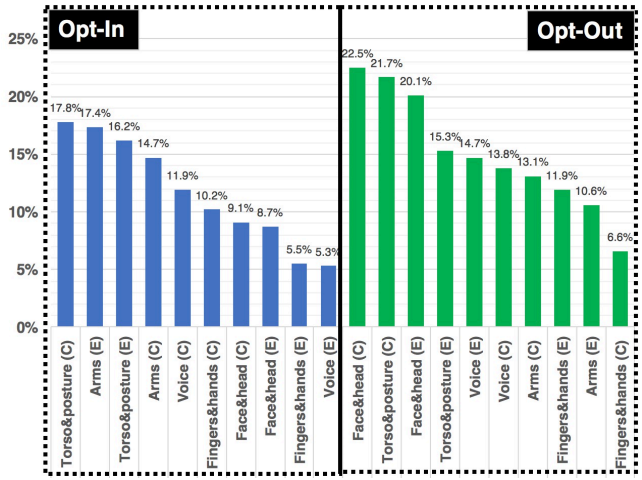

Figure 3. Agreement scores. Columns in blue represent 'opting-in' tasks while 'opting-out' tasks are in green.

Agreement Scores

We calculated agreement scores to find out more about consensus among participants regarding their suggestions. For this, we use Vatavu and Wobbrock's equation [33] which gives insight into which of the referents received the most uniform proposals. If all the proposed gestures are identical the agreement score is 100%, while solely unique proposals result in 0% agreement. The equation for calculating the agreement scores is [33]:

$$\mathcal{AR}(r) = \frac{|P|}{|P|-1} \sum_{P_i \subseteq P} \left(\frac{|P_i|}{|P|} \right)^2 - \frac{1}{|P|-1}$$

where r is the referent in the set of referents R, P is the set of proposals for the referent r, and P_i is a subset of identical symbols from P [33]. This equation extends [40] by adding two correcting factors depending on the number of participants and number of elicited proposals (see [33] for details). To calculate the scores we used the Agreement Analysis Toolkit (AGATe v2.0) [33].

The obtained agreement scores for this study range from 5.3% to 22.5% (Figure 3). While the results lie close together it is noticeable that the first 7 top-ranked gestures are occupied by all four torso & posture tasks and twice by face & head challenges. While 'opting-out' using face and head ranked 1st and 3rd the agreement scores for these body parts are in the lower quarter for 'opting-in'. And though the overall agreement scores are relatively low, in the remainder of the paper we will discuss patterns we observed in the kinds of gestures that were proposed.

Most Mentioned Gestures vs. Selected Favorites

Next, we wanted to investigate the relationship between most mentioned gestures and the ones selected as preferred gestures. To verify this, the groups were asked to choose their favorite gesture which was then compared with the most mentioned proposals. To be 'most mentioned' a gesture had to be named by at least two or more pairs. If there was a tie among the most mentioned gestures, and a tie among the most frequently favored proposals, the ones that happen to match were considered for the analysis presented in Table 2. The results shown in the table illustrate that with only two exceptions (marked in orange) the most often suggested gestures did in fact match with the nominated favorites.

However, not every pair that proposed a frequent gesture chose it as their favorite as well. For the commercial scenario 'opting-out with voice' received the lowest congruence (8 mentions vs. 2 favorites), while the other referent categories are more consistent. Regarding the entertainment scenario, we report that in four cases the gesture proposals were too diverse to identify matches. The pairs either proposed a different favorite each (see 'opting-in' for entertainment scenario) or a gesture was not named by enough people (see 'opting-out' for entertainment scenario). The reason for this might lie in the participants' own aspiration of coming up with something more unique and exceptional: "*I think I would get a little bit more silly with this one (...) because you're at a fun place*" (P8).

Scenario	Referent		Most mentioned gesture	Most frequent favorite
COM	Opt-in	Face & head	Smile (5)	Smile (3)
		Fingers & hands	Waving (5)	Waving (4)
		Arms	Swipe (4)	Swipe (4)
		Torso & posture	Standing in front of it (8)	Standing in front of it (3)
		Voice	Hello (8)	Hello (6)
	Opt-out	Face & head	Shake head (7)	Shake head (4)
		Fingers & hands	Close fist (2)	Close fist (2)
		Arms	Make 'X' (6)	Pull down arm (6)
		Torso & posture	Turning away (8)	Turning away (4)
		Voice	Goodbye (8)	Goodbye (2)
ENT	Opt-in	Face & head	Move head from side to side (4)	–
		Fingers & hands	Jazz hands/spirit fingers (both 3)	–
		Arms	Big wave (3)	Big wave (3)
		Torso & posture	Twist hips (4)	Twist upper body (3)
		Voice	Greeting (5)	Greeting (4)
	Opt-out	Face & head	Shake head (6)	Shake head (2)
		Fingers & hands	Closing fist (4)	Closing fist (2)
		Arms	–	–
		Torso & posture	Turn away (5)	Turn away (3)
		Voice	Goodbye (8)	Goodbye (6)

Table 2. Comparison of most often named gestures with most frequent favorites (numbers in brackets are frequency).

Demonstrated Gestures: Recurring Interaction Patterns

In the following section, we will describe five interaction patterns that we observed during the study and after transcription and thematic analysis of the video material from the elicitation sessions.

Agreement	Disagreement
Yes (e.g., nodding)	No (e.g., shaking head)
Good (e.g., thumbs up)	Bad (e.g., thumbs down)
Begin (e.g., waving)	Stop (e.g., crossing arms)

Table 3. Semantic mappings of agreement and disagreement in order to opt-in or opt out.

Interaction Pattern 1: Expressing Agreement or Disagreement

The scenario that was introduced to the participants was about enabling them to be more explicit about their decision to take part in, finish or refuse an interaction with a public installation and hence avoid frustration (e.g. due to a failing start of an interaction) or invasion of their private space (e.g. accidental capturing). We noticed that many proposed gestures that would express their *agreement or disagreement*. Most of the times, a pair's starting and ending gestures would follow that same theme. The theme of agreement can be further divided into sub-themes. Once they are labelled, they resemble the verbal attempt to pursue the same goal (see Table 3). Gestures corresponding to these sub-groups could be allocated based on the participants' comments. Good/Bad does in fact only contain one gesture, which is putting a thumb up or down, but it occurred multiple times among the participating pairs. Opposed to that, the sub-themes Yes/No and Stop/Do comprise a set of different gesticulations, which are, for example, nodding (Yes), shaking the head (No), stretching out the palm of the hand or crossing the arms like the letter X in front of the body (both representing Stop).

Goal	Gesture
Open	• Open a fist (P8, P13) • Separate touching palms in a linear movement (P1) • Circular arm gestures, described with key word • Voice command: *"Start"* and *"Begin"* (P2)
Close	• Close fingers to make a fist (P3, P6, P8, P12) • Close eyes for a few seconds (P6, P10, P11) • Lowering one or both arms from a higher position (P7, P10) • Voice command: *"End"*, *"Finish"*, *"Stop"* (P5, P8, P16)

Table 4. Example gestures representing open and close as synonym to start and finish and interaction.

Interaction Pattern 2: Open and Close Gestures

Opting-in and opting-out has also been interpreted as 'opening' and 'closing' an interaction which resulted in participants proposing gestures to show they wanted to physically open or close a session with the installation. While performing these gestures, participants explained: *"This is something like open"* (P1*), "An open kind of thing"* (P12), "Like closing the shades" (P9) and *"Something like opening and closing the window"* (P4). The resulting gestures were then a metaphorical representation of the verbs open and close. Table 4 shows examples of how this was achieved. The voice commands mentioned in Table 4 are not literal expressions of the goal, however they share the same idea of defining the beginning and end of an interaction. Only something that has been opened is ready to start responding while an ending command closes the system.

Interaction Pattern 3: Greeting vs. Command

In the voice/speech-only interaction method, the participants came up with two forms of communication. To address the public installation, they either used a command or a greeting. Four pairs came up with both types before deciding on their favorite. Overall, greetings were eventually the most frequent expressions chosen to be the consensual submission of the group. This was the case for both scenarios (commercial and entertainment), as well as both situations (opting-in and opting-out). The greetings included usual expressions such as *"Hello"*, *"Hi"*, *"Bye"*, or *"Goodbye"*. Whereas commands were either action words like *"Exit"* (P12) or requests such as *"Show me something"* (P9) or *"Interact with me"* (P13).

Reasons why the pairs eventually decided on greetings for the mirror scenario might be because they had the feeling of speaking to some sort of virtual assistant (e.g. *"I kinda associate it with turning something on to complete a task"* P9) or due to their habit of how to address existing systems (*"...it's simple as in Hey Siri"* P15). Regarding the entertainment scenario, there was in fact a slight shift towards more informal expressions observable, such as *"What's up"* (P13) or *"Hey, yo"* (P4).

Furthermore, nonverbal sounds were suggested as alternatives. This can inspire concepts of public interfaces that are set in international environments with a multilingual user population, such as airports. The participants' ideas included, for example, whistling to opt-in, and blowing, or the sound *„Shh"* to opt-out.

So far, the presented patterns appear to share a certain similarity. The participants often chose opposing concepts to fulfil the cycle of opting-in and opting-out from an interaction. While this might have been an intuitive result among some groups, we also observed that other pairs did prefer this logical construct on purpose. They reflected which criteria a gesture should meet before coming up with one: *„(...) opposite of what we did to turn it on"* (P9) and also confirmed that *„(...) how ever you started it, it makes sense to finish it. So if you said hello or whatever you say goodbye"* (P16).

Overall, the word 'opposite' fell several times as the participants tried to describe and justify the gestures they were suggesting. For instance, P5 tried to explain a gesture after showing it: *"Yeah, like the opposite of the whole opening thing"*. Although all groups knew before the study started that each scenario would involve two situations (opting-in and opting-out), intentions to build opposing pairs were not observable until reaching the stage of finding gestures to end interactions.

Interaction Pattern 4: Humanization and Anthropomorphism
This pattern summarizes the observations that participants explicitly mentioned an anthropomorphized interaction with the installation. For example, P1 states *"I like the idea of treating it like a human. I prefer a greeting over a 'start'"* when speaking about the interactive mirror.

The proposals for voice categories that were discussed earlier in this paper show how greetings and natural questions were suggested apart from computer-like commands (e.g. 'exit', 'turn on', etc.) to communicate. These greetings sometimes 'imitated' an interpersonal approach by adding a personal reference to it (e.g. *"Hello, mirror"* – P8). P15 explains that the mirror could have *"a name [on the screen] that says* 'Say Hello, [name]'*"*.

This direct, personal addressing of the installation as a form of personal communication was also observed with the proposed body gesture set. Waving, bowing, and even blowing a kiss, can also be understood as transfer of familiar greeting or farewell procedures.

Interaction Pattern 5: Legacy Bias
In our study, we also observed that the presented scenarios evoked certain expectations among the participants that were derived from their previous experience with other technology solutions. Hence, they proposed gestures or interaction procedures that resembled those that are already part of their daily life. This so called legacy bias is often reported in other studies [13,23,28] and can sometimes become a challenge for the exploration of novel interaction approaches as participants hold on to the known and familiar, P16 says: *"It makes sense if it replicates systems you already use like phones and laptops and it has generic swiping gestures or back buttons so that people can intuitively interact with it"*.

A recurring theme participants displayed was suggesting WIMP concepts for the installation such as start or exit buttons on the screen that could be 'pressed' by using the arm, finger, or palm as mouse cursor. Also, Apple's iOS principles are among the findings. ‚Slide to unlock' known from iPhones was mentioned as well as ‚Hello Siri' when suggestions for a voice command were discussed and justified. With '*OK, Mirror'* another existing speech recognition assistant was used as model since this command resembles Google's *'OK, Google'*. That the participants were in fact imprinted by touchscreen technologies became clear during the freestyle category when they brainstormed gestures without any given limitations. Remarks such as *"Naturally I would just touch it"* (P5) were frequent. Furthermore, for some proposed gestures the similarity to touch interactions was seen as advantage, for instance 'swiping' is *"good because we are used to iPads, touchscreens and stuff"* (P4).

Findings about Appropriateness of Gestures in Public
This section presents the result of the questionnaire about the appropriateness of the proposed gestures in the two scenarios commercial/entertainment (the questionnaire was handed out after the elicitation tasks were completed).

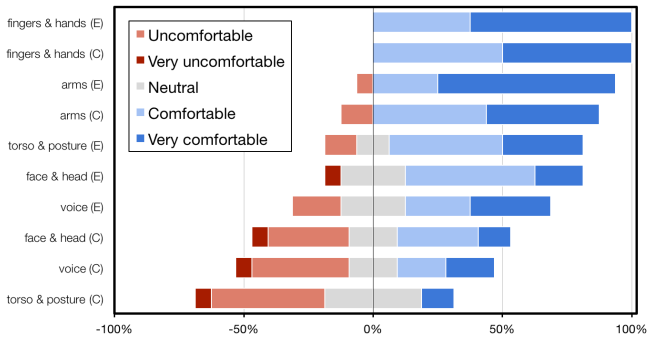

Figure 4. Participants' answers to the questions "How comfortable (=not embarrassed) would you be interacting through gestures in public?".

Comfort level for gesture performance in public spaces
After both scenarios were explored, the participants were asked to assess their level of comfort using different body gestures (i.e., whether or not they would be embarrassed to perform the gestures in public). In analyzing the responses, we noticed differences showing that the participants felt generally more comfortable in an entertainment setting for interactions through gestures (see Figure 4). The results show comfort levels for gestures with hands as the most comfortable in both scenarios (with arms the second most comfortable), and it is notable that the 'voice' and 'torso' category ranked much lower in the participants' preference (see Figure 5). The small variation between hands and arm gestures within a non-commercial setting matches the comments of the participants stating that they would not mind using their arms to produce bigger gestures while they prefer smaller and more discrete movements when they are in a store. There was overall a low standard derivation for these question amongst participants: SD<1.20 for 8 out of 10 of the questions.

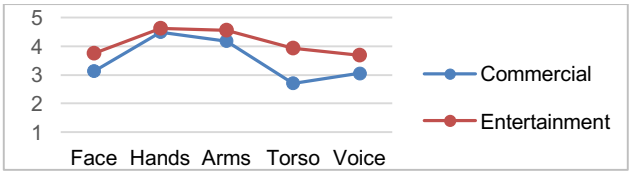

Figure 5. Comparing means of comfort levels (from 1= very uncomfortable, to 5=very comfortable).

For torso gestures in a more playful context the mean of comfort levels came third with a value very close to 4 which represents 'comfortable'. Torso gestures for commercial installations ranked only 2.69, thus the least favorite form of interaction. Regarding face and voice interactions the participants had similar opinions about them: thinking of a commercial situation they were rather "undecided" about their comfort, but if asked about the other scenario their statement was closer to 'comfortable' (face=3.75, voice=3.69).

Less Preferred Situations for Gesture-based Control

The second part of the questionnaire involved a free text question reading: *Can you think of situations, places, certain set-ups where you would not want to interact through gestures?* The aim of this question was to learn more about other factors that could influence the rejection of installations in public spaces. The participants answered by naming characteristics an undesirable situation/place/set-up would fulfil (e.g. somber) or in form of enumerations of concrete places (e.g. hospital).

We clustered the responses to identify key themes. Participants would not like to interact through gestures in a place/situation/set-up where one or more of the following descriptions apply: *crowded, quiet, formal,* or *serious*. While the pattern of 'crowded' was found in 37.5% of the answers, the others became apparent in 18.75% of the answers. That crowded and busy places are seen as inconvenient for gestures can be due the participants' perception that gestures need a lot of space or alternatively because they are frequented by more people that *"(...) can easily spot you"* (P15).

It is particularly noteworthy that 'library' was mentioned by 31.25% of the participants, either describing a quiet environment (n=3) or a serious setting (n=1) although gestures do not really produce sounds (expect voice commands) nor are they socially embarrassing (see previous section). For 'serious' interactions, participants gave concrete examples: P6 would not like to interact through gestures at a *"sensitive exhibition like WW2"* which means the topic the installation deals with is important, too. Furthermore, public services were mentioned such as parliament, GP, embassy, etc. (P7, P11).

There were also practical remarks: One participant (P14) addressed the necessity of being physically available to perform gestures which is why *"*(….) *places where I'd be holding a lot of stuff like heavy bags"* seem inconvenient situations to expect arm or hand gestures. Such could happen, for example, in a supermarket which was also mentioned by another participant (P16).

Participants' vision for implementation

We also asked participants where they would approve gesture-based interaction in public environments: *Can you think of a scenario/context where gesture control could be useful/beneficial?* This question was answered in form of lists of places or by describing certain circumstances. The analysis consisted again of manually clustering and defining recurrent themes.

Over 50% of the participants' answers were about entertainment applications. They can be summed up in P7's words as *"lively, artful, fun places"*. Most often parks (n=4), museums (n=4), and festivals (n=3) were mentioned. Two participants also considered amusement parks. Shopping situations were only mentioned three times. This result coincides with the high levels of comfort people expressed for the entertainment scenario. Gestures are seen as a playful form of interaction that fits best in recreational environments. Furthermore, ambient noise was an aspect that 25% of the participants considered. They either mentioned that gesture control could be useful in *"situations where I have to be quiet"* (P14) but also in loud places (n=3). This seems to be the logical consequence considering that quiet spaces where previously named as undesirable for voice interaction. Additionally, 12.5% of the participants can image that gesture control could be beneficial for *"education purposes"* (P4), e.g. in installations that help to *"visualize ideas or concepts"* (P8) in classrooms or shared workspaces.

	Statement
1	If I saw a gesture installation in a public space I would be curious and approach it.
2	I would not mind if people were watching me using a gesture installation **outdoors** like in a park.
3	I would not mind if people were watching me using a gesture installation in a **busy** urban space (e.g. square, shopping promenade).
4	If an installation is a form of **art**, I would not mind being captured by its camera even if I wasn't interacting with the installation directly.
5	If an installation is a form of **art**, I would not mind being captured by its camera from a further distance even if I had not realized its presence.
6	If an installation is for **entertainment** purposes, I would not mind being captured by its camera even if I wasn't interacting with the installation directly.
7	Performing gestures makes me tired.
8	I would not mind if people were watching me using a gesture installation in a store.
9	If an installation is for **entertainment** purposes, I would not mind being captured by its camera from a further distance even if I had not realized its presence.
10	I find gesture control physically uncomfortable
11	If an installation is for **advertisement** purposes, I would not mind being captured by its camera even if I wasn't interacting with it directly.
12	If an installation is for **advertisement** purposes, I would not mind being captured by its camera from a further distance even if I had not realized its presence.

Table 5. Statements presented to participants.

General perceptions on the use of public installations

In the last section of the questionnaire participants were asked to rate their level of agreement for a set of given statements (see Table 5).

Figure 6 presents the outcome with the most prominent result being the high likelihood the participants would approach and perhaps use interactive systems in public spaces. Most participants answered that they do not find gestures physical uncomfortable, though it was an almost even split for the question if the gestures make them tired. They also do not mind if many people observe them during an interaction, in either a busy place (M=3.44) or quieter outdoor spaces (M=4.06). However, being watched in a store is perceived more critical (M= 2.88).

Furthermore, the questionnaire asked about people's opinion on being captured by a body tracking camera. Their acceptance varied depending on the purpose of the installation: most positive are art installations, then entertainment (e.g. game), and lowest for advertisements. Their acceptance was slightly higher if they were aware of the presence of a camera installation. The results for art and entertainment installations are more consistent, while partici-

pants felt very strongly about advertisement. The reason for this result might lie in the social value these areas enjoy in general. While advertisement might be seen more negatively and might evoke privacy concerns when user data is collected (e.g. in form of images in this case), arts and entertainment might be perceived as something that can lead to a positive experience.

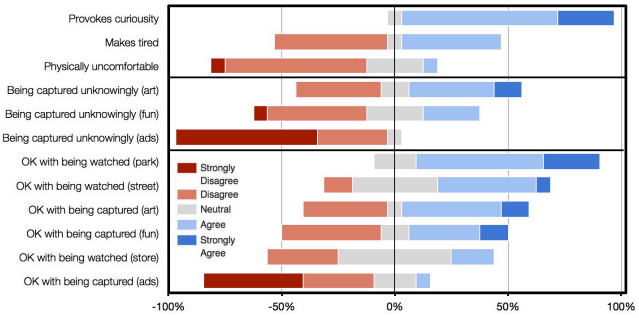

Figure 6. Participants' agreement with presented statements.

DISCUSSION

Cultural Influence and Other Factors

The multi-cultural background of the participants enriched this study giving it a broader perspective on communicating through gestures. We were repeatedly reminded about the cultural differences regarding the meaning of gestures. This phenomenon is studied in-depth in other disciplines such as sociology, communication studies, or psychology [3]. Mauney et al. [24] did make cultural similarities and differences the focus of their research on user-defined gestures for touch interfaces. Related to those findings, during the elicitation study participants made us aware of the affirmative meaning shaking one's head has in some parts of India (contrary to western cultures), and that quickly tossing the head back can mean 'No' in Balkan countries and Greece.

Cultural differences not only exist regarding the meaning of gestures but also as to the significance gestures have for the communication accuracy of a language/culture [13]. This means that some cultures are more likely to complement and 'support' their speaking with gestures. Hence, people might use gestures more frequently and are therefore more open to gesticulate in general. Both are important aspects to consider when designing and evaluating installations with gesture-based input, or when designing future experiments. Furthermore, future studies could also help to identify the impact of other factors (for example, the age of participants, or their previous experience with gesture systems) on the set of proposed gestures.

Social Context and Location

Our results align with earlier research, such as Harrison & Dourish's [15] and Akpan et al.'s [1], identifying that 'place' (social context) is determinant for the ability of an installation to encourage interaction and facilitate engagement with potential users. Our study indicated that the favored interaction technique of users (e.g. with hands or face) also depends on the social context of an installation. The more playful an interaction is, the more comfortable our participants were with using uncommon body gestures to interact with it. The same applies to their acceptance of body-tracking technology. Future studies could help finding out more about people's perception of how acceptable and comfortable such gestures would be in different real-world environments.

Communicating Interactivity and Possible Gestures

The interactive systems we are designing also need to address the challenge of how they can communicate the possible gestures to a person interacting with the display (Bellotti et al. describe this as '*Attention: How do I know the system is ready and attending to my actions?* [4]). Systems would need to provide feedback that they allow interaction, how to interact, and which options are available. This also applies to opt-in and opt-out gestures, which could be communicated with graphical and text explanations on the screen. Such feedback could, for example, built on strategies proposed by Müller et al., where they found mirrored and silhouette visualizations very effective for engaging passers-by in interaction [28].

Unintentional Interactions

Last, there is the challenge of how to minimize unintentional interactions [4]: for example, a person might greet another person by waving their hand, which the system – accidentally – interprets as opt-in. Out of the gestures found in our study, this would apply in particular to the anthropomorphic and conversational gestures (which are more likely to happen in every day conversations with other people), but might be less critical for other gestures (such as crossing the arms in front of the face or doing circular arm movements). As discussed before, ambiguity and unintentional interactions could be possibly addressed with visual or auditory feedback by the system. Furthermore, a person does need ways to recover quickly from mistakes (e.g., how to end an interaction directly when opting-in by mistake).

CONCLUSION

In this paper, we described a gesture elicitation study that was conducted to investigate how people can use gestures to opt-in or opt-out from interactions in public environments. We provide a comprehensive list of gestures which were suggested by the participant of this study. Our analysis revealed interaction patterns indicating that the gesture proposals are often nonverbal representations of mental models associated with 'opting-in' and 'opting-out'. Our insights also include what type of contexts are preferred for interactions with gestures as well as which body parts people feel most comfortable to use for these gestures. In case of 'opting-in' and 'opting-out' it would be desirable to develop a universal gesture vocabulary, with the goal to allow passers-by to reject an undesired interaction as quickly as possible. In this regard, our identified gestures and patterns are a starting point for finding appropriate forms for people opt-in or opt-out from interactions with future public installations.

REFERENCES

1. Akpan, I. et al. 2013. Exploring the effects of space and place on engagement with an interactive installation. *Proc. of CHI '13*. ACM.
2. Alexander, J. et al. 2012. Putting your best foot forward: investigating real-world mappings for foot-based gestures. *Proc. of CHI '12*. ACM.
3. Archer, D. 1997. Unspoken diversity: Cultural differences in gestures. *Qualitative Sociology*. 20, 1.
4. Bellotti, V. et al. 2002. Making sense of sensing systems: five questions for designers and researchers. *Proc. of CHI '02*. ACM.
5. Boring, S. et al. 2014. The Dark Patterns of Proxemic Sensing. *Computer*. 47, 8, 56–60.
6. Brignull, H. and Rogers, Y. 2003. Enticing people to interact with large public displays in public spaces. *Proc. of Interact '13*.
7. Chan, E. et al. 2016. User Elicitation on Single-hand Microgestures. *Proc. of CHI '16*. ACM.
8. Chong, M.K. and Gellersen, H. 2011. How users associate wireless devices. *Proc. of CHI '11*. ACM.
9. Connell, S. et al. 2013. A Wizard-of-Oz elicitation study examining child-defined gestures with a whole-body interface. *Proc. of IDC '13*. ACM.
10. Dalton, N.S. et al. 2015. Display Blindness?: Looking Again at the Visibility of Situated Displays using Eye-tracking. *Proc. of CHI '15*. ACM.
11. Fischer, P.T. and Hornecker, E. 2012. Urban HCI: spatial aspects in the design of shared encounters for media facades. *Proc. of CHI '12*. ACM.
12. Gatto, I. and Pittarello, F. 2012. Prototyping a gestural interface for selecting and buying goods in a public environment. *Proc. of AVI '12*. ACM.
13. Graham, J.A. and Argyle, M. 1975. A cross-cultural study of the communication of extra-verbal meaning by gestures (1). *Int. Journal of Psyc*. 10, 1, 57–67.
14. Greenberg, S. et al. 2014. Dark Patterns in Proxemic Interactions: A Critical Perspective. *DIS '14*. ACM.
15. Harrison, S. and Dourish, P. 1996. Re-place-ing space: the roles of place and space in collaborative systems. *Proc. of CSCW '96*. ACM.
16. Holmes, T.G. 2007. Eco-visualization: combining art and technology to reduce energy consumption. *Proc. of C&C '07*. ACM.
17. Huang, E.M. et al. 2008. Overcoming assumptions and uncovering practices: When does the public really look at public displays? *Proc. of Pervasive '08*. Springer.
18. Ju, W. et al. 2008. Range: exploring implicit interaction through electronic whiteboard design. *Proc. of CSCW '08*. ACM.
19. Kinect for Windows. *http://www.microsoft.com/ en-us/kinectforwindows/develop/*.
20. Kinect for Windows Retail Clothing Scenario Video: *https://www.youtube.com/watch?v=Mr71jrkzWq8*.
21. Kinect-Powered Cube Unveiled at Decibel Festival: *https://www.youtube.com/watch?v=1jzHcXZFJMw*.
22. Krogh, P. et al. 2004. "Help Me Pull That Cursor" A Collaborative Interactive Floor Enhancing Community Interaction. *Australasian Journal of Inf. Systems*. 11, 2.
23. Mäkelä, V. et al. 2014. Information wall: evaluation of a gesture-controlled public display. *Mobiquitous '14*.
24. Mauney, D. et al. 2010. Cultural similarities and differences in user-defined gestures for touchscreen user interfaces. *Proc. of CHI EA '10*. ACM.
25. Morris, M.R. et al. 2014. Reducing legacy bias in gesture elicitation studies. *interactions*. 21, 3, 40–45.
26. Morris, M.R. 2012. Web on the wall: insights from a multimodal interaction elicitation study. *ITS '12*.
27. Müller, J. et al. 2009. Display blindness: The effect of expectations on attention towards digital signage. *Proc. of Pervasive '09*. Springer.
28. Müller, J. et al. 2012. Looking Glass: A Field Study on Noticing Interactivity of a Shop Window. *CHI '12*.
29. O'Hara, K. et al. 2008. Understanding collective play in an urban screen game. *Proc. of CSCW '08*. ACM.
30. Perry, M. et al. 2010. WaveWindow: public, performative gestural interaction. *Proc. of ITS '10*. ACM.
31. Preece, J. et al. 2015. *Interaction Design: Beyond Human-Computer Interaction*. John Wiley & Sons.
32. Seyed, T. et al. 2012. Eliciting Usable Gestures for Multi-display Environments. *Proc. of ITS '12*. ACM.
33. Vatavu, R.-D. and Wobbrock, J.O. 2015. Formalizing agreement analysis for elicitation studies: new measures, significance test, and toolkit. *CHI '15*.
34. Vatavu, R.-D. and Zaiti, I.-A. 2014. Leap gestures for TV: insights from an elicitation study. *TVX '14*. ACM.
35. Verma, S. et al. 2015. A Study for Investigating Suitable Gesture Based Selection for Gestural User Interfaces. *Proc. of IndiaHCI '15*. ACM.
36. Vermeulen, J. et al. 2015. Proxemic Flow: Dynamic Peripheral Floor Visualizations for Revealing and Mediating Large Surface Interactions. *INTERACT 2015*.
37. Vogel, D. and Balakrishnan, R. 2004. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. *Proc. of UIST '04*. ACM.
38. Walter, R. et al. 2013. StrikeAPose: revealing mid-air gestures on public displays. *Proc. of CHI '13*. ACM.
39. Wang, M. et al. 2012. A Public Advertising Display that Captures and Preserves the Attention of a Passerby. *Proc. of PerDis '12*. ACM.
40. Wobbrock, J.O. et al. 2005. Maximizing the guessability of symbolic input. *Proc. of CHI EA '05*. ACM.
41. Wobbrock, J.O. et al. 2009. User-defined gestures for surface computing. *Proc. of CHI '09*. ACM.