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

Isabel Benavente Rodríguez
HCI-E MSc Final Project Report 2016
UCL Interaction Centre, University College London
Supervisor: Dr Nicolai Marquardt

ABSTRACT
Compared to interfaces with conventional input devices (e.g. touchscreen) where beginning and end of an interaction is easy to manipulate by the users, an interactive system with a gesture-capturing camera often lacks to allow passers-by this choice as its body tracking is continuously running. This paper explores how people could use gestures to explicitly express their wish to commence an interaction, reject, or finish it. For this, a gesture elicitation study with 16 participants (grouped in pairs) was conducted to explore gestural and voice commands to opt-in or opt-out from interactions with installations in public settings. The participants were introduced to two scenarios (commercial scenario & entertainment scenario). They proposed 430 gestures to opt-in or opt-out from an interaction, and we calculated agreement scores that indicate comparatively high consensus for torso gestures in general as well as for 'opting-out' with face and head only. Furthermore, patterns were identified showing that participants aimed for non-verbal representations of the opposing pairs ‘close & open’ and ‘agreement & disagreement’. It became also evident that greetings were chosen over commands which relates to the pattern of ‘humanizing’ the installation. From questionnaires we learnt that participants preferred gestures with hands and arms for both scenarios, while their overall acceptance for other body parts was higher for the entertainment setting than for the commercial scenario. Regarding topics such as public comfort, public appropriateness, and privacy concerns, we report that again the scenario is the determinant factor. The participants rated installations that were a piece of art or entertainment as less intrusive and discomforting than if it was for commercial or advertisement purposes. To conclude, design implications are discussed which could support the user-centred design of future public installation.

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

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

MSc Contribution Type
Interaction Techniques.

1. INTRODUCTION
The affordability of body motion-tracking devices made it possible to develop technology installations for public spaces that can be operated through gestures (e.g. large interactive screen). However, this still novel interaction technique can be a challenge as people might have little experience with gesture-based commands.

So far, currently more common touch screen interfaces have shaped people’s expectations and habits regarding its use. While instructions on the screen such as ‘touch here to start’ or exit buttons help to clearly determine start and length of an interaction with the system, it can be difficult for users to verify their status (e.g. ‘Have I been noticed?’) when communicating through gestures.

Furthermore, once the willingness to engage in an interaction exists it can be frustrating and lead to the dropout of the users if they have difficulties “connecting” with the system. It is likewise unpleasant to get involved involuntarily with it only because the single steps of the interaction procedure blend into each other.

In fact, this form of dysfunctional interaction might be (in some more playful cases) intentional by the designers because of its comical potential. Nevertheless, it leaves curious users disappointed and with an unsatisfying sense of no control over the interactive system as well as uninterested passers-by feeling their privacy has been invaded.

It is therefore important that the handling of an installations is as comprehensible and intuitive as possible in order to satisfy all parties concerned that use a certain public space.

In this study we will investigate how people can use gestures to clearly opt-in or opt out from interactions with an interactive system in public environments. The aim is to
explore ways that support potential users in being more explicit when communicating their consent or refusal to interact with a body tracking system.

During the gesture elicitation study that was conducted for this purpose following the example of current research [7, 22,30,31], we invited 8 pairs to propose gesture commands representing ‘opting-in’ and ‘opting-out’ using successively 4 different parts of their body and also their voice. In combination with two scenarios that were introduced (commercial vs. entertainment setting), this resulted in a list of 20 referents.

We present the results of the elicitation study together with calculations regarding the agreement among participants for each referent and discuss how the underlying logic of their suggestions which let emerge patterns could eventually shape the future interaction vocabulary for gesture-based systems.

Furthermore, we also address topics related to interactions in public environments such as comfort, public appropriateness, and privacy concerns by evaluating questionnaires that accompanied the elicitation study.

2. LITERATURE REVIEW
The following section gives an overview of literature relevant to this dissertation project. A discussion of the current state of research on public interfaces sets context for a closer look at gesture-based interactions and how the development of gestures was previously approached. Following this, we discuss the gesture elicitation method as it will be applied in this study.

Technology in Public Spaces
Today, interactive technologies are no longer only findable in homes and workplaces but in public environments, too. There is no limit regarding the variety of purposes and functions an installation can offer. This and their diverse looks (ranging from large displays, to interactive floors [17], to art installations visualising real-time data [14]) can make them an exciting object within a public space.

While some installations are rather task-oriented like WaveWindow [26], which allows to inspect delicate objects for example in museums, or Gatto et al.’s [10] prototype to buy goods, others such as Information Wall [18] were designed to fulfil locally relevant information needs.

That colourful public displays can be an eye-catching brand ambassador has been tested by companies such as LEGO [37] or Pepsi [34]. With interactive game displays in shop windows or on vending machines, both companies availed themselves of the curiosity this novel technology may provoke.

Studying Technologies in Public Spaces
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 us and our environment. The insights are intended to contribute to the improvement of their design by addressing i.a. the visibility of interfaces, the likelihood of engaging potential users, or their capability of holding a user’s attention.

A public space can offer a large number of visual stimuli for people to process at once. It is likely that their glances will wander around trying to cover as much as possible. Whether they notice and decide to interact with an interactive display in a public setting can depend on different factors.

Müller et al. [24] argue that when it comes to public displays passers-by show a certain blindness for them. Their research revealed how people filter out their presence based on the expectation of being exposed to irrelevant content. The study was later reviewed by Dalton et al. [8] using eye-tracking resulting in new insights that passers-by apparently do look at public displays more than previously assumed, but only briefly and as long they are still far away. Both studies were carried out in a different context and time, however it becomes clear that public interfaces are not natural stop-and-stare eye catchers but are perceived rather peripherally.

A ‘successful’ installation will therefore need to be able to compete for attention. For this, Huang et al. [5] discourage from relying on the large size of the interface or installation as only eye-catcher but to take advantage of the environment and make use of already existing objects in the vicinity that could draw additional attention if an interactive system is deployed nearby. To find these objects an analysis of the spatial arrangement of the area can be beneficial to discover the movement and flow of people within the public space as well as the directions they are likely to walk and hence look at.

A favourable position for public installations can furthermore be explored consulting Fischer & Hornecker [9] which gives guidance on how a space around an installation can be divided into different sub-spaces whose existence and size have varied impact on a deployment. For instance, a generous activation space (area from where an installation can be seen and trigger curiosity) or social interaction space (where people have the change to gather and observe other users interacting) can promote the visibility as well as the engagement of passers-by. This can also lead to a honey pot effect where the presence of interested people creates a sociable buzz that further attracts more spectators [5].

Spatial aspects are not only important for an installation to be noticeable but also for the participation of potential users. Starting from the common definition of space and place, Akpan et al. [1] prove in their research the superiority of latter, representing the social context that is shaped with time through people’s memories, experiences,
norms, etc., over ‘space’ being the physical platform where certain behaviour is facilitated or not.

More precisely, although a space fulfils superficial requirements (e.g. visible and large), people are less likely to interact with a system installed there if its purpose and functions mismatch their associations or perception of appropriateness.

For example, it is likely that visitors of a theme park are more open to games on a large display since fun and playfulness is something to expect there. However, there can still be park areas where people are less comfortable to engage in a game such as situations and space contexts where they feel too exposed or vulnerable. For instance, in places where their association is ‘being watched’ (e.g. in front of a large seating area that resembles an audience).

Exposure or social embarrassment is also described in [5] as potential barrier for participation. In order to keep it as low as possible Brignull & Rogers [5] suggest to design the installation in a way that it is able to equip quickly and easily potential users with the kind of information they need to evaluate how beneficial/enjoyable a participation might be. Their scepticism will be reduced if they know exactly what to expect when walking up. Hence, observing the installation in action should answer how to use the system, what the consequences are, how long an interaction is, whether it is comfortable, etc. In the best case, this also becomes clear when standing further away (as additional attraction factor). In the context of gesture-based interactions the design should therefore avoid to involve gestures that could be perceived as embarrassing or foolish.

**Interaction type for interactive installations: air-based gestures**

Today, technology has advanced such that it is able to recognize people’s body and limb gestures through motion-sensitive cameras and other computer sensing and vision techniques [27]. They have been implemented in commercially available applications that can be connected to video game consoles such as Sony’s EyeToy for the PlayStation or Microsoft’s Kinect for the Xbox. The cameras track the movements of the players and project them onto the screen by translating them into the movements of an avatar or a tool. This natural input method facilitates the access for users of a wider array of ages and gaming skills with little or none experience in handling controllers [27].

But the affordability and hardware compatibility of the Kinect system moreover opened up whole new possibilities for alternative concepts other than gaming in home environments. It has not only found a use in open source projects of a large tinkerer community [35,36] but also in prototyping for academic research. Technology-based marketing campaigns or art installations too implement Kinect to add an interactive dimension to their deployment.

Taking advantage of the touch-less interaction a body motion camera facilitates, O’Hara et al. [25] developed an interactive system that allows surgeons to handle it during operations while keeping their hands sterile. Instead of using a (potentially germ-infested) keyboard and mouse the surgeon can navigate through X-Rays and CT images using hand gestures.

For a similar kind of workspace scenarios where the hands are not available or occupied Alexander et al. [2] investigated how foot gestures can function as command input and developed design recommendations emphasizing the advantages of rate-based techniques over displacement based techniques.

Regarding a public environment, touch-less interaction can be an option for cases where direct contact with a physical interface is not possible. Reasons for this might include:

- The dimensions of the installation require a certain distance for the optimal visual angle
- Remote control is not viable in order to reduce supervision costs of the equipment
- The technology of an installation is too delicate to be touched by the large amount of passers-by a public space usually experiences
- Physical constraints of limiting users

**Understanding and Designing Gestures**

According to [27] one of the biggest challenges when designing air-based gesture input is “to consider how a computer system recognizes and delineates the user’s gestures”. 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.

A way to help make this happen is by giving the user clearer instructions on how exactly to perform certain interactions. A more precise gesture execution might support the system in better recognizing. Furthermore, it helps first-time users to familiarize with the system’s operation.

Based on this idea, Walter et al. [32] 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, or in other words: the action of opting-in.

Their research resulted in a proposal of three strategies: spatial division (the gesture is constantly shown in a certain screen area), temporal division (the application is interrupted from time to time to reveal the gesture in between), and integration (gesture hints are embedded in the application). They furthermore proposed the “Teapot” as initial gesture [32].
Another strategy to ease air-based interaction for users is to set the focus on developing gestures that are easy to understand, relate, and execute.

**Gesture Elicitation Studies**

In order 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 centre of gesture design and use them as source for interaction recommendation. This technique shows participants the desired result of a command (called referent) and asks them to suggest a gesture (called symbol) that would trigger it [23].

With a certain flexibility that allows to include additional parameters based on specific research interests, the method of eliciting gestures might run as follows:

In advance to the study referents are framed which represent actions a system is able to perform. For instance, a system could have a menu for users to scroll through. ‘Scrolling’ is here the command for which researchers want to explore new input techniques away from familiar devices with arrow buttons. In the study, user/participants are then confronted with the execution and the result of the command (e.g. through a Wizard of Oz study setup [22]) to better visualise the actual interaction goal. With this goal in mind, the objects are invited to brainstorm gestures how they would bring about the desired result.

However, regarding this method there is certain consensus of researchers of the field that legacy bias is a big challenge and limitation when designing an elicitation study [13,23,28]. Legacy bias can occur when participants (unintentionally) build upon their experience with other technology during the process of proposing interactions. The outcome might therefore be based on existing habits and familiarity with previous action-reaction concepts rather than on true intuition and usability.

In recent years, elicitation studies have been applied to investigate different home scenarios. While [30] investigated how gesture control can be used for controlling TV systems and [22] explored opportunities to use internet browser function on a TV device through gestures, [7] and [30] chose gaming as object of study. They all observed patterns among their participants’ gesture proposals that can be consulted for the development of future systems.

If the explored gestures are supposed to be implemented in a design concept, metrics can be used to compare collected data in order to “create a single canonical gesture set” [23]. Such metrics can be, for example, acceptance of a gesture, recall likeness, or agreement among participants.

Ideally, the gestures participants proposed for a referent exhibit some overlaps as this suggests that certain proposals could be likewise plausible to a broader group and hence user friendly interactions. Wobbrock et al.’s [33] formula helps to calculate the match level by putting frequently occurring gestures in relation to the sum of all suggested gestures. The specialty of the equation is that it is able to consider parallel emerging patterns and hence express a more meaningful agreement score.

To advance the systematization of elicitation analysis Vatavu & Wobbrock [29] developed a toolkit that expands the agreement term by introducing measures that also capture disagreement rates and coagreement rates. The toolkit also includes a statistical significance test that allows to compare agreement rates.

Drawing on the example of previous elicitation studies and analysis techniques, we extend earlier research by applying said method for gestural interaction in public spaces. In particular, we will explore if the body parts used to perform certain gestures (that start, refuse or abandon an interaction) matter when public embarrassment and social acceptance become variables. We will furthermore investigate if gestural interaction and for this purpose inevitable body tracking through a camera impacts the disposition of people to engage with an interactive system outside their homes. Both will contribute to the state of knowledge on factors that can influence the likelihood of an installation being accepted and used. Additionally, relating to the previously introduced concept of place and space our study compares two environmental contexts in order to review the suggestion of different results.

### 3. GESTURE ELICITATION STUDY FOR OPTING-IN AND OPTING-OUT OF INTERACTIONS WITH PUBLIC DISPLAYS

This section describes how the study of this paper was designed and carried out. It draws on the methodologies that were introduced in the preceding paragraphs.

**Participants**

The participants of this study were recruited from the circle of acquaintances of the experimenter. They undertook the study in pairs and were either partners or friends.

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).

Based on their self-reported daily use of different electronic devices (smartphone, tablet, e-reader, laptop, desktop computer, gaming console), we assume a high technology proficiency: fifteen of the participants use a smartphone and laptop on a daily basis and eight use at least three devices (smartphone, tablet, laptop) every day.

With only one exception all of the 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 at one point of their lives.

To compensate for their time, each participant had the chance to enter a raffle for one of four £25 Amazon vouchers.

**Study Design**
Following the example of [30] and [22] an elicitation study was designed to explore gestures with participants. In order 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 in [22] this creates a relaxed environment where the tasks given become stimuli to brainstorm as a team rather than a challenge demanding instant answers.

The groups of two were presented sequentially 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 their refusal/abandonment. The first scenario took place in a commercial setting while the installation of the second scenario was for entertainment and recreational purposes.

Later on, restrictions were given allowing the participants to only use certain body parts for proposing gestures (see Figure 3). These consisted of using:

- Face & head only
- Fingers & hands only
- Arms only
- Torso/posture only
- Voice as command

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 at least two. To move on to the next category, the team had to agree on a favourite and declare it consensus of the group. ‘Agree to disagree’ was also a valid submission.

In order to facilitate an accurate motion analysis, the entire sessions were video-recorded allowing the collection of qualitative data. 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 free text questions investigating ideas on benefits and limitations of gesture-based installations in public spaces.

**Refinements of Study Design through Pilot Study**
Before running the main experiment, a pilot study was conducted to review the consistency and procedure of the study. The weaknesses detected were corrected, however no second round of pilot testing took place. This was at the experimenter’s discretion who assessed the changes as minor. Furthermore, the results of both questionnaire and gesture elicitation tasks were not considered for the evaluation as they originated under different parameters.

The initial draft of the study design intended to test half of the pairs with a reversed order of the scenarios in order to include an additional element of randomization. The pilot confirmed the previously emerged hypothesis that participants would be more familiar with a commercial context and might be over challenged and confused by the unusual challenge of brainstorming body gestures if they had to start within a less explored entertainment setting. To avoid that confusion overshadowed the first half of the study, the randomized scenario order was waived allowing the participants to ‘warm-up’ with a familiar scenario before facing a more abstract challenge.

Overall, the wording of the verbally given explanations was revised to guarantee a more precise description of the tasks. Therefore, examples were prepared explaining why the intention to explicitly opt-in or opt-out could exists in the first place:

- to terminate the stand-by mode of an installation (opt-in)
- to communicate disagreement of being captured (opt-out)
- to make sure confidential data is erased from the interface after an interaction (opt-out).

Moreover, the instructions stated on the questionnaire were also refined. For instance, the version handed out in the pilot asked the participants about their level of comfort when performing gestures with certain body parts. It was shown that “comfort” offered too many interpretations such as the absence of physical effort while the experimenter’s aim was to investigate whether the participants would be embarrassed having to use certain body parts in public.

**Materials and Referents**
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.

At the beginning of the session the pairs received an information sheet explaining the outline of the study and implications regarding data collection and privacy. Informed consent was obtained.

The study started with the participants filling out a pre-questionnaire collecting demographic data and experience levels. The study concluded with a follow-up questionnaire. The videos used to demonstrate the scenarios are both intellectual property of Microsoft and are published on YouTube under usernames ‘Microsoft’ or ‘KinectforWindows’ [16,20]. The participants were shown edited 30-second versions with highlight sequences (see Figure 1 and 2).

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.

**Procedure**

To ensure all crucial information from the introduction sheet was clear to the participants, the experimenter repeated key facts and ethics of the project and gave the objects the opportunity to clarify questions. Then, the 3 parts of the study were explained: pre-questionnaire, elicitation, follow-up questionnaire.

After filling out the first questionnaire which collected demographic data and assessed the experience level with technology in general and more specifically gesture-based interactions, the participants were shown the first scenario video.

In it, an interactive mirror in a clothing store was presented that gave its users a preview of how different garments would fit them. The user used 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 where gesture-based installations can be used to carry out a certain marketing agenda, influence purchase decisions, or place personalized advertisement.

In the debrief of the video the experimenter addressed how the camera of an installation like this would always capture potential users, for example by ‘dressing up’ passers-by with random items such as hats and glasses in order to catch their attention.

The participants were invited to imagine they would like to have more control over the start and end of a potential interaction. In other words, they want their consent to be more explicit when it comes to interacting with an installation like the mirror.

To help their imagination in order to brainstorm together, the pairs were presented with a representation of the mirror installation that occasionally gave visual feedback, such as a flashing ‘Hello’ or ‘Bye’. This was adopted as point of reference by the participants to address their gestures to.

The elicitation started with the instruction to find gestures that would represent ‘opting-in to an interaction’. While the first attempt of the participants was spontaneous and ‘freestyle’, they were later challenged to find at least two gestures under certain limitations (e.g. face and head only).

For each limiting category as well as the freestyle category, the pairs had to agree on their favourite 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 [22] in order to obtain very basic indications to weight the proposed gestures against each other. When all limitation categories were discussed the same procedure was repeated for ‘opting-out from an interaction’, once again starting with a freestyle warm-up.
Then, for the second scenario, the structure of the elicitation study was similar. It was also introduced by a video that showed a cube (approx. 3x3 metres) which imitated the dance rhythm of a user through colourful palpitating animations on its frosted acrylic glass surface.

The cube represents the sort of playful public installation whose purpose it is to entertain its users and incite their creativity. Some might only exist for the sake of art; some can have a function.

Following the video, the experimenter mentioned the disposition of entertainment installations to capture people in the surroundings with its camera at all time.

Placed in front of the screen with a representation of the cube that occasionally changed colours and animations, the pair commenced the familiar gesture proposal procedure for gestures that would represent ‘opting-in’ and ‘opting-out’ in this scenario.

The sessions which lasted between 45 and 60 minutes concluded with a post-study questionnaire where the participants rated on Likert scales their level of comfort (= not embarrassed) when performing gestures with the different body parts that were explored through the given limitations. Finally, a post-study questionnaire asked the participants to evaluate their experience and comfort with air-based gestures which were then analysed quantitatively.

4. FINDINGS
This section presents the findings of the elicitation study as well as the insights from the questionnaires. The results of the elicitation study can be divided into quantitative and qualitative results. The former begins with an overview of the 430 proposed gestures and shows how these are distributed over the body part categories. Then agreement scores will highlight which of these categories prompted the most congruent gestures among the participants. Additionally, we will also review how the favourites nominated by each group of participants relate to their overall mention 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. Some body part restrictions prompted more suggestions than others which is illustrated in Figure 4. Except for ‘Opt-in with arms in an entertainment setting’ all body categories received the asked minimum of two gestures or more per pair. Overall, ‘arms’ received the least suggestions in three out of four opt-in/opt-out situations, while ‘voice’ prompted the most suggestions in total (across all situations).
<table>
<thead>
<tr>
<th>Sc.</th>
<th>Referent</th>
<th>Proposed gesture</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>face&amp;head</td>
<td>Nodding, smile, long blink, move head from side to side, wink, draw a circle with the head, bend head, stick tongue out, shake head, raise eyebrows, own body with eyes, make long eye contact, nod up (once), kiss, move head back and forth (= cultural gesture: ‘The Egyptian’ dance move), turn head to one side</td>
</tr>
<tr>
<td>Opt</td>
<td>fingers&amp;arms</td>
<td>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</td>
</tr>
<tr>
<td>-in</td>
<td>torso&amp;posture</td>
<td>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)</td>
</tr>
<tr>
<td>C</td>
<td>voice</td>
<td>‘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</td>
</tr>
<tr>
<td></td>
<td>face&amp;head</td>
<td>Shaking head (No), long blink, turning head round and round, nod, bow, chin up, blow</td>
</tr>
<tr>
<td>Opt</td>
<td>fingers&amp;arms</td>
<td>Closing fist, press button wit hands as cursor, show palm (cultural gesture: Stop), hand flick, gesture: show fingers 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</td>
</tr>
<tr>
<td>-out</td>
<td>arms</td>
<td>Pulling lifted arm(s) down, make a cross (‘X’) with both arms in front of body, gesture: show fingers 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</td>
</tr>
<tr>
<td>torso&amp;posture</td>
<td>Turning away and showing back for a few seconds, crouch/slouch forward, bow, turn half away (45°), squat, shrug, shake “it off”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>face&amp;head</td>
<td>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)</td>
</tr>
<tr>
<td>Opt</td>
<td>fingers&amp;arms</td>
<td>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</td>
</tr>
<tr>
<td>-in</td>
<td>arms</td>
<td>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)</td>
</tr>
<tr>
<td>torso&amp;posture</td>
<td>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)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>face&amp;head</td>
<td>Shaking head (No), head down for a few seconds, scrunch face, blow kiss, nod up (once), long blink, wink, look up with entire head</td>
</tr>
<tr>
<td>Opt</td>
<td>fingers&amp;arms</td>
<td>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</td>
</tr>
<tr>
<td>-out</td>
<td>arms</td>
<td>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</td>
</tr>
<tr>
<td>torso&amp;posture</td>
<td>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</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Complete list of all distinct gestures proposed by the participants.
Agreement Scores
So far, we presented a complete set of the elicited gestures. It can gain in significance if it is complemented by data about the consensus among participants regarding their gesture suggestions. For this, Wobrock et al. (2005) introduced a formula to calculate agreement rates which ultimately 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. Wobrock et al. [33] explain their equation as follows:

\[
A = \frac{\sum_{r \in R} \sum_{P \in P_r} \left( \frac{|P|}{|P_r|} \right)^2}{|R|} \cdot 100
\]

The obtained agreement scores for this study range from 7.5% to 25.0%. They are shown in Figure 5. While the results lie rather close together it is indeed noticeable that the first 7 spots are occupied by all four torso & posture tasks and twice by face & head challenges. While ‘opting-out’ using face and head only ranked 1\textsuperscript{st} and 3\textsuperscript{rd} it is interesting to see that the agreement scores for these body parts are in the lower quarter for ‘opting-in’.

Most mentioned gestures vs. appointed favourites
This section is based on the hypothesis that gestures which were mentioned several times among the pairs might be particularly intuitive. However, only because many people had the same idea for a gesture this does not necessarily mean they actually describe it as their favourite interaction method. Consequently, the question arises whether being intuitive does influence the likelihood of being preferred over other gestures. To verify this, the groups were asked to choose their favourite gesture which was then compared with the most mentioned proposals. To be ‘most mentioned’ a gesture had to be named by at least 2 or more pairs. If there was a tie among the most mentioned gestures, and a tie among the most frequent favoured 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 favourites.

However, not every pair that proposed a frequent gesture chose it as their favourite as well. For the commercial scenario ‘opting-out’ with voice received the lowest congruence (8 mentions vs. 2 favourites), while the other referent categories are rather 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 favourite 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).

<table>
<thead>
<tr>
<th>Sc.</th>
<th>Referent</th>
<th>Most mentioned gesture</th>
<th>Most frequent favourite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opt-in</td>
<td>face&amp;head</td>
<td>Smile (5)</td>
<td>Smile (3)</td>
</tr>
<tr>
<td></td>
<td>fingers&amp;arms</td>
<td>Waving (5)</td>
<td>Waving (4)</td>
</tr>
<tr>
<td></td>
<td>arms</td>
<td>Swipe (4)</td>
<td>Swipe (4)</td>
</tr>
<tr>
<td></td>
<td>torso&amp;posture</td>
<td>Standing in front of it (8)</td>
<td>Standing in front of it (3)</td>
</tr>
<tr>
<td></td>
<td>voice</td>
<td>Hello (8)</td>
<td>Hello (6)</td>
</tr>
<tr>
<td>Opt-out</td>
<td>face&amp;head</td>
<td>Shake head (7)</td>
<td>Shake head (4)</td>
</tr>
<tr>
<td></td>
<td>fingers&amp;arms</td>
<td>Close fist (2)</td>
<td>Close fist (2)</td>
</tr>
<tr>
<td></td>
<td>arms</td>
<td>Make ‘X’ (6)</td>
<td>Pull down arm(s) (6)</td>
</tr>
<tr>
<td></td>
<td>torso&amp;posture</td>
<td>Turning away (8)</td>
<td>Turning away (4)</td>
</tr>
</tbody>
</table>
In, finish or refuse an interaction with them to be more explicit about their disposition to take part in a public installation. The broader purpose behind this action: it is about enabling accessible, the for opting in and opting out has also been interpreted as starting and ending an interaction which resulted in participants proposing gestures to show they wanted to open or close a session with the installation.

The explanations the participant gave were essential in order to understand the gestures as they included very specific mentioning of key words. Thus, quotes included:

- “This is something like open” (P1)
- “An open kind of thing” (P12)
- “Kinda close the curtain” (P7)
- “Like closing the shades” (P9)
- “Something like opening and closing the window” (P4)

Moreover, the command “End” was proposed by 2 pairs (P1&P2, P15&16) in the voice category which in fact is quite explicit. The resulting gestures were then a metaphorical representation of the verbs open and close. Table 4 shows examples how this was achieved.

<table>
<thead>
<tr>
<th>Sc.</th>
<th>Referent</th>
<th>Most mentioned gesture</th>
<th>Most frequent favourite</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>voice</td>
<td>Goodbye (8)</td>
<td>Goodbye (2)</td>
</tr>
<tr>
<td></td>
<td>face&amp;head</td>
<td>Move head from side to side (4)</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>fingers&amp;arms</td>
<td>Jazz hands/spirit fingers (both 3)</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>arms</td>
<td>Big wave (3)</td>
<td>Big wave (3)</td>
</tr>
<tr>
<td></td>
<td>torso&amp;posture</td>
<td>Twist hips (4)</td>
<td>Twist upper body (3)</td>
</tr>
<tr>
<td></td>
<td>voice</td>
<td>Greeting (5)</td>
<td>Greeting (4)</td>
</tr>
<tr>
<td></td>
<td>face&amp;head</td>
<td>Shake head (6)</td>
<td>Shake head (2)</td>
</tr>
<tr>
<td></td>
<td>fingers&amp;arms</td>
<td>Closing fist (4)</td>
<td>Closing fist (2)</td>
</tr>
<tr>
<td></td>
<td>arms</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>torso&amp;posture</td>
<td>Turn away (5)</td>
<td>Turn away (3)</td>
</tr>
<tr>
<td></td>
<td>voice</td>
<td>Goodbye (8)</td>
<td>Goodbye (6)</td>
</tr>
</tbody>
</table>

Table 2. Comparison of most often named gestures with most frequent favourites. Numbers in brackets refer to the frequency.

**Demonstrated Gestures: Recurring Interaction Patterns**

The transcription and thematic analysis of the video material from the elicitation sessions showed that participants ‘created’ gestures intending to nonverbally represent a certain semantic model. In their reasoning patterns emerged that will be explained in the following sections. They were elaborated using an adaption of Braun & Clarke’s (2006) approach of thematic analysis. Since the gestures are often not self-speaking, the data set also includes the explanations the participants gave during the demonstration as this is a reliable source to learn about their intentions and chain of thought behind a gesture.

**Interaction Pattern 1: Expressing agreement or disagreement**

To make the challenge of finding gesture representations for opting-in and opting-out from an interaction more accessible, the experimenter explained to the participants the broader purpose behind this action: it is about enabling them to be more explicit about their disposition 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).

Many participants were then inclined to focus on finding gestures that would express their agreement or disagreement. Most of the times, a pair’s starting and ending gestures would follow that same theme. However, this was not imperative.

This 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).

**Table 3. Semantic representation of agreement and disagreement in order to opt-in or opt out. Arranged in opposing pairs if possible.**

<table>
<thead>
<tr>
<th>Agreement</th>
<th>Disagreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Good</td>
<td>Bad</td>
</tr>
<tr>
<td>–</td>
<td>Stop</td>
</tr>
</tbody>
</table>

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 due to its high occurrence among the participating pairs, it was undeniable to create an own sub-group.

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).

**Interaction Pattern 2: Close and open**

Opting in and opting out has also been interpreted as starting and ending an interaction which resulted in participants proposing gestures to show they wanted to open or close a session with the installation.

The explanations the participant gave were essential in order to understand the gestures as they included very specific mentioning of key words. Thus, quotes included:

- “This is something like open” (P1)
- “An open kind of thing” (P12)
- “Kinda close the curtain” (P7)
- “Like closing the shades” (P9)
- “Something like opening and closing the window” (P4)

Table 4 shows examples how this was achieved.
Informal expressions observable, such as associating it with turning something on, speaking to the mirror scenario, e.g., “Hi”. The greetings included usual expressions from traditional greetings like “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 cube scenario, there was in fact a slight shift towards more informal expressions observable, such as “What’s up” or “Hey, yo”. However, opting-out did not involve more creative approaches than traditional greetings like “Bye”.

Furthermore, worth to mention are the suggestions that did not involve language. Nonverbal sounds could be a solution to overcome the obstacle users that are not proficient in the local language might face. This can inspire conceptions of public interfaces that are set in international environments with a multilingual user population, such as airports. The participants’ ideas included, for example, for example, whistling to opt-in, and blowing or the sound „Shh“ to opt-out.

An interim conclusion: the opposite logic
So far, the presented patterns appear to share a certain similarity. In particular, 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 the stage of finding gestures to end interactions would begin.

Interaction Pattern 3: Greeting vs. command
The limiting category that allowed to use voice only as interaction method showed that 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 favourite.

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 cube scenario, there was in fact a slight shift towards more informal expressions observable, such as “What’s up” or “Hey, yo”. However, opting-out did not involve more creative approaches than traditional greetings like “Bye”.

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

<table>
<thead>
<tr>
<th>Goal</th>
<th>Gesture</th>
</tr>
</thead>
</table>
| Open  | • Open a fist (P8, P13)  
|  | • Separate touching palms in a linear movement (P1)  
|  | • Circular arm gestures that were 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 down (P7, P10)  
|  | • Voice command: “End”, “Finish”, or “Stop” (P5, P8, P16) |
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 hand, 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 as 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).

**Insights about Appropriateness of Gestures in Public**

This section presents the result of the questionnaire that was handed out after the elicitation tasks.

**Comfort level for gesture performance in public spaces**

The participants were asked to assess their level of comfort using different body gestures (as in not embarrassed to perform them in public) after both scenarios were explored. This gave them the opportunity to compare both experiences and rate them in relation to each other.

In the analysis clear differences were observable illustrating that the participants felt generally more comfortable in an entertainment setting for interactions through gestures (see Figure 6).

As the comfort levels for gestures with hands are the highest in both scenarios, it can be assumed that this is the body part the participant would prefer over others if they had the choice. 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 if in a store.

The low standard derivation of the collected quantitative data suggests agreement amongst the participants. It was not over 1.00 for 6 out of 10 Likert scales while SD≥1.20 occurred only twice. This allows to consider means as tool to compare collected comfort levels (see Figure 7).

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 favourite 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).

![Figure 6. Participants’ answers to the questions “How comfortable (=not embarrassed) would you be interacting through gestures in public?”.](image-url)
Situations of no interest in 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 apart from the form of interaction (e.g. arm gestures) 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. sombre) or in form of enumerations of concrete places (e.g. hospital). All of the 16 objects gave an answer.

The rather small size of the data set allowed a straightforward manual clustering of same or similar entries using Excel. By doing so, key themes emerged.

According to them, the 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
- serious

While the pattern of ‘crowded’ was findable 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). This provides an indication that they might mind being observed (see Literature Review).

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). To understand this observation in-the wild validations an of gesture-based installation placed there can provide more insight.

Except for the comment “(...) where everyone is in suit and tie” (P3) which explains what ‘formal’ might mean we do not have information what a place like this implies for gesture-based interactions.

Much different is the pattern ‘serious’. Here the participants gave more 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). An interpretation could be that systems with gesture interaction are seen as less efficient to meet user goals in said contexts or alternatively this could be related to privacy concerns as systems in these places might deal with sensitive user information.

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

Within the same questionnaire segment a further free text question explored the potential areas where gesture-based interaction in public environments would find a higher approval. The question read: Can you think of a scenario/context where gesture control could be useful/beneficial?

This question too was answered in form of lists of places or by describing certain circumstances. The analysis consisted again of manually clustering and defining recurrent themes.

The most distinct finding regarding this questionnaire part is that the answers of 50% of the participants can be associated with an entertainment setting. 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. Compared to this, shopping situations were part of the answer only three times.

This result coincides with the high levels of comfort people expressed for the entertainment scenario. Apparently, gestures are seen as a more playful form of interaction that fits best in recreational environments. Unfortunately, this questionnaire could not gather more information about the reasons for this notion. However, contextual inquiries or interviews could complement these findings.

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. But yet, it remains unresolved whether this is because noisy is associated with a “lively” (P1) scenery.

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 “visualise ideas or concepts” (P8) in classrooms or shared workspaces.

**General perceptions on the use of public installations**

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

<table>
<thead>
<tr>
<th></th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>If I saw a gesture installation in a public space I would be curious and approach it.</td>
</tr>
<tr>
<td>2</td>
<td>I would not mind if people were watching me using a gesture installation outdoors like in a park.</td>
</tr>
<tr>
<td>3</td>
<td>I would not mind if people were watching me using a gesture installation in a busy urban space (e.g. square, shopping promenade).</td>
</tr>
<tr>
<td>4</td>
<td>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.</td>
</tr>
<tr>
<td>5</td>
<td>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.</td>
</tr>
<tr>
<td>6</td>
<td>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.</td>
</tr>
<tr>
<td>7</td>
<td>Performing gestures makes me tired.</td>
</tr>
<tr>
<td>8</td>
<td>I would not mind if people were watching me using a gesture installation in a store.</td>
</tr>
<tr>
<td>9</td>
<td>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.</td>
</tr>
<tr>
<td>10</td>
<td>I find gesture control physically uncomfortable</td>
</tr>
<tr>
<td>11</td>
<td>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.</td>
</tr>
<tr>
<td>12</td>
<td>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.</td>
</tr>
</tbody>
</table>

**Table 5. Statements presented to participants.**

Figure 8 presents the outcome with the most prominent result being the high likelihood the participants would approach and perhaps use interactive systems in public spaces.

Their experience with gesture interaction appears to be rather uncomplicated as most of them said they do not find gestures tiring nor physical uncomfortable. They also do not mind if a lot of 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 a bit more sceptical (M=2.88). Overall, this is a promising result for the future of gesture interaction in public spaces.

Furthermore, the questionnaire asked about their opinion on being captured by a body tracking camera. Their acceptance decreased depending on the purpose of the installation: art, entertainment (e.g. game), advertisement.

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 (M=3) while participants felt very strongly about advertisement.
The reason for this result might lie in the social value these areas enjoy in general. While advertisement often experiences a negative connotation that evokes privacy concerns when user data (e.g., in form of images in this case) is collected, arts and entertainment are perceived as much more ‘harmless’ or even able to return something positive.

5. DISCUSSION
As described in the preceding section, this study produced unambiguous results contributing to the solution of finding ways to opt-in and opt-out from interactions with public displays. We report that there was high consensus among our participants regarding their proposals for gestures with their torso and through posture only. This is quite notable as participants predominantly stated communicating through torso/posture was rather embarrassing. They clearly preferred gestures with hands and arms. Regarding the other body parts, we found out that their acceptance was higher if they were used within a more playful context. Said context was also responsible for the perception of a public installation. While commercial deployments and advertising displays are perceived as intrusive and discomforting, art or game installations enjoy a much better reputation.

The analysis of patterns among the gesture proposals showed that there were tendencies to humanize the installation, for example by using greetings instead of machine commands. Furthermore, the participants demonstrated intentions to find gesture representations for ‘close & open’ and ‘agreement & disagreement’ where we found cultural differences in the execution.

Based on all these findings, in this discussion we will have a closer look at the influence context and culture appear to have on gesture-based interaction with public installations. The section will close addressing design implications and potential directions this topic could take in the future.

Cultural cues
The multi-cultural background of the participants enriched this study giving it a broader perspective on communicating through gestures. It reminded us that there are indeed cultural differences regarding the meaning of gestures. This phenomenon is rather studied in other disciplines such as sociology, communication studies, or psychology [3]. However, Mauney et al. [19] did make cultural similarities and differences the focus of their research on user-defined gestures for touch interfaces.

In conversation with the participants of this elicitation study they made us aware of the affirmative meaning shaking one’s head has in some parts of India (contrary to western cultures). We also learnt that quickly tossing the head back can mean ‘No’ in Balkan countries and Greece. Many more of these cultural distinctions are already widely known. But because semantic representations of gestures are so deeply anchored in our mental models, it is not too difficult to underestimate them.

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 [11]. This means that some cultures are more likely to complement and ‘support’ their speaking with gestures. Hence, we might deduce that they use gestures more frequently and are therefore more open to gesticulate in general. Both are

![Figure 8. Participants’ agreement with presented statements.](image-url)
important aspects to consider when designing and evaluating installations with gesture-based input.

**Contextual cues**
In consequence of our study findings we want to emphasize once more the impact use contexts have on (gesture-based) interactions in public and the general perception of public installations. Our results confirm Harrison & Dourish’s [12] and Akpan et al.’s [1] claim that ‘place’ (social context) is determinant for the ability of an installation to encourage interaction and facilitate engagement with potential users.

We extend this previous work by contributing insight that also the favoured interaction technique of users (e.g. with hands or face) depends on the social context of an installation. The more playful said is, the less uncomfortable are potential users using uncommon body gestures to control it. The same applies to their acceptance of body-tracking technology.

**Design Implications**
We refrain of issuing our table comprising all the proposals from the elicitation study as set gestures to opt-in or opt-out from interactions with future public displays. Much more, we want to point out for the future that it will appear more intuitive to users if they are able to recognise a concept behind the gestures the installation proposes that they can actually associate with the desired interaction result. For instance, affirmative actions should not be represented by gesticulations commonly (culturally dependent) understood as negation.

In case of ‘opting-in’ and ‘opting-out’ it would be indeed desirable to develop a universal gesture vocabulary (for that reason alone to allow passers-by to reject an undesired interaction as quickly as possible). Our paper nevertheless, reveals tendencies how users interpret ‘opting-in’ and ‘opting-out’ in order to translate them into nonverbal symbols. Hence, the interaction patterns ‘close & open’ and ‘agreement & disagreement’ are our contribution to this topic and are a starting point for further exploration.

Furthermore worth to mention in this section is how a certain sensitivity with respect to the caginess of users towards tracking cameras will benefit the design of a novel installation. Large public interfaces are still unusual for an important share of potential users and investigating how to make sure they are not put off will be crucial in this ‘settling-in period’ that we are still in.

**Future Work**
A next step for this study could be to implement the gesture suggestions of this paper in prototypes of public installations that are under development. As part of in-the-wild studies, this would allow to test with a broader spectrum of subjects the appropriateness of said gestures for different scenarios and contexts in order to verify our previous claims. It would furthermore revise the significance of the discussed agreement scores. This reiterative testing procedure would contribute to the validation of our collected gesture proposals.

**6. 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. This project has achieved its aim of contributing insightful results to this topic. We provide a comprehensive list of example gestures which were suggested by the participant of this study. Its analysis revealed interaction patterns indicating that the gesture proposals are often times nonverbal representations of mental models associated with ‘opting-in’ and ‘opting-out’.

Using an established technique of the field, we calculated agreement scores for the proposals obtained during the elicitation study which express the consensus among participants for gestures executed by different body parts. 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. We can also contribute results regarding people’s perception of body-tracking technology in public spaces.

**ACKNOWLEDGMENTS**
I am sincerely grateful to my supervisor Dr Nicolai Marquardt for his exceptional guidance, valuable feedback, patience and encouragements throughout the course of this project. I also thank Dr Steven Houben for taking the time to give me advice, the participants of this study, who shared so much of their creativity with me as well as my fellow HCI-E peers for their friendship and enriching discussions.

Furthermore, I wish to express my deepest gratitude to my family: to my parents for always supporting me and giving me so many opportunities to pursue my interests, and to Tefi & Tobi for being role models. Last but not least, a special thanks to Andy for always offering his help and making sure I am well.

**REFERENCES**


APPENDIX 1: INFORMATION SHEETS, INFORMED CONSENT FORMS, PAPER QUESTIONNAIRES

Attached to this paper are the information sheets and consent forms that were presented to the participants of this study as well as the questionnaires that collected demographic information and quantitative data.
Information Sheet for Participants

You will be given a copy of this information sheet.

1. Title of the research project
Air-based gestures for public interfaces: how to opt-in & opt-out from interactions

This study has been approved by the UCL Research Ethics Committee as Project ID Number: UCLIC/1213/015

2. Contact Details of the study conductor
Isabel Benavente Rodríguez
isabel.rodriguez.15@ucl.ac.uk
+49 176 706 355 36

UCL Interaction Centre
8th Floor Malet Place Engineering building Gower Street
London WC1E 6BT

I would like to invite you to participate in this research project. You should only participate if you want to; choosing not to take part will not disadvantage you in any way. Before you decide whether you want to take part, please read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or you would like more information.

3. Details of Study
• The purpose of the study is to find out more about how people can use gestures to interact with interactive systems in public environments.
• After filling out a questionnaire about previous experiences with installations in public spaces the participant will be asked to brainstorm gestures together with a partner. The session closes with a follow-up questionnaire.
• The session will last approx. 45 minutes.
• There is no harm or risk for the participant when agreeing to take part in this study.

It is up to you to decide whether or not to take part. If you choose not to participate, you won’t incur any penalties or lose any benefits to which you might have been entitled. However, if you do decide to take part, you will be given this information sheet to keep and asked to sign a consent form. Even after agreeing to take part, you can still withdraw at any time and without giving a reason. All data will be collected and stored in accordance with the Data Protection Act 1998.
Consent Form

Title of Project: Air-based gestures for public interfaces: how to opt-in & opt-out from interactions

This study has been approved by the UCL Research Ethics Committee as Project ID Number: UCLIC/1213/015.

Participant’s Statement

I ……………………………………………………………………………………………. (Name)

agree that I have

• read the information sheet and/or the project has been explained to me orally;
• had the opportunity to ask questions and discuss the study; and
• received satisfactory answers to all my questions or have been advised of an individual to contact for answers to pertinent questions about the research and my rights as a participant and whom to contact in the event of a research-related injury.

Furthermore,

• I understand that I am free to withdraw from the study without penalty if I so wish, and I consent to the processing of my personal information for the purposes of this study only and that it will not be used for any other purpose. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.
• I understand that my participation will be taped/video recorded, and I am aware of, and consent to, any use you intend to make of the recordings after the end of the project.
• I understand that the result of my participation will be published as a report. Confidentiality and anonymity will be maintained, and it will not be possible to identify me from any publications.

Signature:                                  Date:

Investigator’s Statement

I ……………………………………………………………………………………………. (Name) confirm that I have carefully explained the purpose of the study to the participant and outlined any reasonably foreseeable risks or benefits (where applicable).

Signature:                                  Date:
Before we start...

a few questions about yourself:

1. Age ...... Occupation ........................................

   Gender  □ male  □ female  □ prefer not to say

2. Have you ever used an installation in a public space that you could operate with your gestures or body movement?
   (e.g. an interactive screen with ads in a mall, an art installation in a museum that reacted to your movements…)

   □ yes  □ no

3. Have you ever played a video game with a camera that tracked your body movements?
   (e.g. Nintendo Wii, Xbox Kinect, PlayStation Move, etc.)

   □ yes  □ no

4. What type of electronic devices do you use in your everyday life?

   □ Smartphone
   □ Tablet
   □ E-reader
   □ Laptop
   □ Desktop Computer
   □ Gaming console
1. How comfortable (=not embarrassed) would you be interacting through gestures (as you brainstormed them today) in **public**? *Please tick the box.*

   **a)** In a commercial setting as seen in the mirror scenario  
   *(e.g. in a store, when looking for or buying things, etc.)*

   **b)** In an entertainment/art setting as seen in the cube scenario  
   *(e.g. in a museum, at a festival, in a park, etc.)*

<table>
<thead>
<tr>
<th>Gestures / Interactions</th>
<th>Very Uncomfortable</th>
<th>Uncomfortable</th>
<th>Undecided</th>
<th>Comfortable</th>
<th>Very Comfortable</th>
</tr>
</thead>
<tbody>
<tr>
<td>...with your <strong>face</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>..with your <strong>hands</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>…with your <strong>arms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>…with your <strong>torso</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>…with your <strong>voice</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. Can you think of **situations, places, certain set-ups** where you would not want to interact through gestures?

   ........................................................................................................

3. Can you think of a scenario/context where gesture control could be **useful/beneficial**?

   ........................................................................................................

4
4. Please indicate your level of agreement for the following statements about gesture-based interactive systems (as the scenarios presented previously):

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Undecided</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>If I saw a gesture installation in a public space I would be curious and approach it.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performing gestures makes me tired.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I find gesture control physically uncomfortable.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I would not mind if people were watching me using a gesture installation in a store.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I would not mind if people were watching me using a gesture installation in a busy urban space (e.g. square, shopping promenade, etc.).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I would not mind if people were watching me using a gesture installation outdoors like in a park.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>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.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>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.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>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.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>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.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If an installation is for advertisement purposes, I would not mind being captured by its camera even if I wasn’t interacting with the installation directly.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>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.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>