



Article

User-Generated Gestures for Voting and Commenting on Immersive Displays in Urban Planning

Guiying Du ^{*}, Auriol Degbelo and Christian Kray

Institute for Geoinformatics, University of Muenster, 48149 Muenster, Germany;
degbelo@uni-muenster.de (A.D.); c.kray@uni-muenster.de (C.K.)

* Correspondence: guiying.du@uni-muenster.de; Tel.: +49-159-024-759-82

Received: 14 April 2019; Accepted: 30 April 2019; Published: 7 May 2019



Abstract: Traditional methods of public consultation offer only limited interactivity with urban planning materials, leading to a restricted engagement of citizens. Public displays and immersive virtual environments have the potential to address this issue, enhance citizen engagement and improve the public consultation process, overall. In this paper, we investigate how people would interact with a large display showing urban planning content. We conducted an elicitation study with a large immersive display, where we asked participants (N = 28) to produce gestures to vote and comment on urban planning material. Our results suggest that the phone interaction modality may be more suitable than the hand interaction modality for voting and commenting on large interactive displays. Our findings may inform the design of interactions for large immersive displays, in particular, those showing urban planning content.

Keywords: public displays; public participation; gestural interaction; mobile interaction; voting; citizen feedback

1. Introduction

Due to the reduction of the price and rapid technological advances [1], public displays have become increasingly prevalent in urban public life. There is an ongoing evolution from ambient non-interactive displays to community displays, and large interactive displays deployed in a variety of real-world scenarios, e.g., shop windows or plazas [2–5]. At the same time, research in exploring large interactive displays has been expanding from addressing technical concerns to studying topics such as public participation [6]. With this shift, one new research direction is emerging which is concerned with the use of large interactive (public) displays to facilitate the consultation of citizens during the urban planning process.

There are three main limitations of traditional public consultation methods (i.e., brochures, leaflets and hearing meetings) when it comes to supporting public participation in urban planning processes (see [7]): lack of interactivity; lack of feeling of immersion; and the lack of a possibility to select specific objects and to comment on them (which is a consequence of the lack of interactivity). Previous research [8–10] pointed out that these limitations led to the loss of the public's interest in urban planning. There is also evidence in the literature [11] that both citizens and local authorities value the possibility of giving/getting feedback through public displays. These findings from previous works have motivated us to examine users' wishes when it comes to using gestures (both hand and phone) to provide their feedback on urban planning content in an Immersive Video Environment (IVE) [12]. In the IVE, we visualized spatial urban planning information using panoramic video footage of the environment overlaid with 3D building models. This type of audiovisual simulation can provide people a feeling of immersion [13], and enhance engagement and learning [14–16].

Designing suitable interfaces for large public displays that support citizens giving their feedback on urban planning designs is still a challenge [17]. In this paper, we focus on exploring two means of interaction, namely hand gestures and mobile devices. Hand gestures may bring a more intuitive way of engaging with immersive digital environments [18–20], where people should stand well away to get the entire urban planning information on the displays. By using hand gestures, people who must sit in the wheelchairs can avoid touching the screen surface and switching between viewing and interacting states to interact with large displays [21]. People, who do not like or have smartphones, can also freely interact with the large displays. Mobile devices, as one of four main interaction modalities (i.e., through touching objects visualized on public displays, through external devices, through tangible objects, and through body movements) with large public displays [18] have become increasingly present in people's everyday lives, with about 59% of the world's population currently owning a smartphone [22].

The underlying scenario of the work is that of a city council wishing to collect citizens' feedback on some alternative designs of early urban planning scenarios. As traditional methods such as public hearings are limited by attendance issues (i.e., basically only those who have time can be informed during the public hearing event), the city council is seeking to complement public hearing events using digital technologies. For this purpose, a public display is deployed in an open space (e.g., in the hall of the city council), and citizens interested can use it to explore the urban planning material, before giving their feedback. They thus can interact with the material at their convenience, one at a time. With this scenario as a backdrop, one goal of the current work is to get a better understanding of the kind of gestures people naturally produce with their smartphones and hands to vote and comment. Another goal is to explore user preferences regarding interaction with public displays using hand gestures or smartphones.

To achieve these goals, we carried out an elicitation study to collect input from users on these interaction modalities in the context of an urban planning scenario. Our study is driven by previous research work [23–25], which resulted in user-defined gestures for different interaction scenarios, i.e., a user-defined gesture set for tabletop systems, a set for navigating a drone, and user-defined gestures to connect mobile phones with tabletops and public displays. Based on our literature review of the research field (i.e., designing interfaces on large public displays that support diverse citizens giving their feedback), this eliciting input from users has not been done yet, which suggests that there is still a gap that needs to be filled. The system we envision is a large immersive display that enables the general public to interactively explore and give their feedback regarding design alternatives during urban planning processes. Example scenarios include engaging the public to give their feedback on designing specific facilities for a neighborhood or a city, i.e., building a new music center in a city; adding a garden house in a community; having a teaching building in a university; to name just a few.

The main contributions of this work are two-fold: (1) a set of hand gestures and smartphone gestures for voting and commenting on large interactive displays. These gestures pave the way for the design of interactive displays with support the move from a lower level (i.e., informing) to a higher (i.e., consultation) on Arnstein's [26] ladder of citizen participation; (2) qualitative user perceptions on the advantages and drawbacks of using hand gestures and smartphones for giving feedback on urban planning material on large interactive displays. These contributions can help in the design of gestural interfaces for (immersive) large displays that are used to facilitate higher levels of participation in urban planning.

2. Related Work

The topic of the paper is situated at the intersection of public display, public participation, gestural interaction, and mobile interaction. This section reviews previous work in these areas. As observed in [17], the words "engagement" and "participation" are sometimes used interchangeably in the literature. In keeping with this practice, work on citizen engagement is also included in this brief review.

2.1. Citizen Engagement in Urban Planning Processes

One important part in public participation theory is the redistribution of power among different stakeholders, e.g., government, institutions, communities, and citizens. Sherry Arnstein's ladder of citizen participation [26] is one of the most popular theories in citizen engagement research. As a typology, her theory suggests how much power citizens get at five different participation levels. Another related public engagement typology was proposed by Rowe and Frewer [27]. This typology identified three public engagement levels, which were defined based on the direction of the information flow and the nature of the information. Since the 1950s empowering citizens in decision-making processes has been intensively proposed with the rapid development of participation theories and models [26,28–32]. As an essential part of decision-making, the spatial urban planning process also calls for citizen participation, such as regional development, transport and mobility, design, and use of open spaces in a city/community. Planners have also published a vast number of theories, models and case studies in recent decades [33,34]. Fostering high-quality information flow among citizens and other stakeholders is one essential part of achieving a successful public participation [27,35].

Citizens conventionally contribute their information in verbal or written form during specific events that are organized by local governments or planning institutions such as public hearing, workshops, and face-to-face dialogues [36,37]. In practice, though it is crucial to integrate citizens, the cost of participation, e.g., the cost of educating oneself about the issue sufficiently, seems to be a severe barrier drawing citizens back from participation [38]. The development of information and communication technologies (ICT) brings about new methods to facilitate public participation in urban planning. Public displays, as one kind of ICT, hold great potentials in citizen participation as introduced in the next subsection.

2.2. Public Displays for Public Participation

The interplay of public displays and public participation has been investigated from various perspectives in the literature. Veenstra et al.'s study [39] brought forth that interactivity is beneficial in engaging citizens, as well as stimulating them to spend more time with, and around a public display. Steinberger et al. [40] also found that interaction was effective in enabling users to vote Yes/No questions through catching people's attention and lowering participation barriers. Hespanol et al. [41] indicated, along with the same lines, that public displays increase participation by encouraging group interaction. As to higher engagement of people, Wouters et al. [42] investigated the potential of citizen-controlled public displays (i.e., displays whose content is controlled by one citizen or household, rather than by a central authority such as a local government or a commercial agency). The authors reported that a more sustained engagement with citizen-controlled public displays can be enforced through a publication process that is explicitly distributed among multiple citizens, or delegated through some open and democratic process. However, Du et al. [17] pointed out that current research has mainly targeted low levels of public participation (e.g., informing citizens) in urban settings, and that work is still needed to explore the use of public displays for higher levels of public participation (e.g., consult citizens or collaborate with them).

Regarding citizens' feedback, Hosio et al. [11] reported that offering the possibility for direct feedback on general urban issues was highly appreciated by citizens in their study. In Valkanova et al.'s research [4], the deployment of Reveal-it! confirmed the ability of urban visualization displays to inspire individual reflection and social discussion on an underlying topic. A comparison of papers, web forms, and public displays in [43] led to the conclusion that public displays are instrumental in generating interest: they manage to attract more comments on each of three different urban topics from people (though feedback on public displays comes at the price of being noisy). Recent relevant research on public displays for urban planning has been mostly focusing on planning support systems, and multi-users settings. For instance, Mahyar et al. [44] designed, implemented and evaluated a Table-Centered Multi-Display Environment, a multi-user platform, to engage different stakeholders in the process of designing neighborhoods and larger urban areas. However, we want to explore the

use of public displays for public consultation in urban planning, in which a single user setting will be needed.

Augmented reality (AR) and virtual reality (VR) technologies are commonly used in urban planning areas for representing urban planning contents [7,45–48]. People found that these technologies have the potential to open new ways for non-professional participation in urban planning. For instance, Dambruch and Krämer [46] presented a web-based application for representing the planned building in 3D models, which aimed to leverage public participation in urban planning. Similarly, Howard et al. [7] implemented a 3D cityscape environment to represent a non-existent city and let people leave their feedback as comments. Besides, van Leeuwen et al. [47] investigated the effectiveness of VR in participatory urban planning by comparing VR headset, VR on the computer, and a 2D plan on papers in citizen participation process concerning the re-design of a public park. Hayek [48] explored the effectiveness of abstract and realistic 3D visualization types as well. However, these ways of technologies also bring some barriers in common: first, additional devices (e.g., personal computers) were needed for people, who wanted to get access to participation; second, the interaction with the urban planning information through computers was either in the hand of experts or not user-centered.

In sum, public displays have great potential in the context of public participation in urban planning. Feedback from citizens on public displays are quite desirable, and AR or VR (including 3D models) are promising technologies for representing urban planning contents for non-professionals. However, more efforts are still needed to explore user-centered interaction with urban planning contents for citizen consultation in the context of urban planning.

2.3. Previous Elicitation Studies

In a gesture-elicitation study, end users are individually prompted with the desired effect of an action/task and asked to propose a gesture that would bring that effect about [49]. Wobbrock et al. [23] presented their work eliciting users' tabletop gestures for interaction design by running the elicitation study. Since then, elicitation studies, became a popular technique in participatory design [50] and various research areas, e.g., free-hand gestures for TV controlling [51], free-hand gestures for uni-modal input in AR [52], surface and motion gestures for mobile interaction [23,53]. Single-hand micro-gestures, which could have application in a wide range of domains, were elicited in [54]. However, as documented in the literature [55,56], the elicitation methodology may also bring several issues. Morris et al. [49] proposed that more research efforts were needed to reduce legacy bias, i.e., the fact that user gesture preferences are influenced by their previous interaction experience with WIMP (windows, icons, menus, and pointing) icons. However, Köpsel et al. [57] held the opinion that instead of rejecting legacy bias, it would be more helpful to take advantage of it. Besides, Tsandilas [56] pointed out that the raw agreement rates [23] should be carefully used depending on the type of study design and user groups.

Overall, we think elicitation studies can provide an effective technique for directly collecting inputs from end users on interaction methods, which would be more intuitive and natural. Therefore, we adopt this method to design user-centered interactions with the urban planning contents on the large public display.

2.4. Gestural Interaction with Public Displays

The use of gestures to interact with public displays has been an active research topic for many years. Inspired by Wobbrock et al.'s [23] seminal work, several researchers further explored user gestures for a variety of tasks and application scenarios. The review of about 65 papers in [58] crystallized the insight that most gestures so far have been designed for selection, navigation, and manipulation tasks.

Panning and zooming with gestures was the topic of [19,59]. Fikkert et al. [59] identified about 27 gestures to manipulate (i.e., pan and zoom) a topographic map on a public display. Nancel et al. [19] observed that linear gestures were generally faster than circular ones for panning and zooming.

They also found that participants generally preferred two-handed techniques over one-handed ones for this task. Large datasets exploration was the subject of [60], and the authors produced a set of gestures to enable interaction with active tokens. Walter et al. [61] investigated the selection of items on public displays and pointed out that by default (i.e., without hint or instruction), people use a combination of pointing and dwelling gesture to select items they want. Rovelo et al. [62] identified a user-defined gesture set for interacting with omnidirectional videos through an elicitation study. Walter et al. [63] found that spatial division (i.e., permanently showing a gesture on a dedicated screen area) was an effective strategy to highlight the possibility for gestural interaction to users, and help them execute an initial gesture called *Teapot*. Ackad et al.'s in-the-wild study [20] showed that the icon-tutorial was effective in supporting learning gestural interactions with hierarchical information on public displays. Overall, this subsection illustrates that much work has been done on user gestures for public displays. These have, however, barely touched on the topic of providing feedback in participation processes which is the core of the current article.

2.5. Mobile Phone Interaction with Public Displays

Using mobile phones to interact with large public displays has become a research focus in recent years [64]. Boring et al. [65] compared three different interaction techniques (Scroll, Tilt and Move) for continuously controlling a pointer on a large public display using a mobile phone. Their experiment revealed that *Move* and *Tilt* can be faster for selection tasks but introducing higher error rates. Lucero et al. [66] designed the MobiComics prototype, which allowed a medium-sized group of people to create, edit, and share comic-strip panels using mobile phones. Using mobile phones, people could also share panels on two large displays. Ruiz et al. [53] proposed a useful distinction between surface gestures (i.e., two dimensional), and motion gestures (i.e., three dimensional), and introduced gestures to perform tasks such as placing/answering a call or switching apps on mobile phones. Liang et al. [67] propose surface and motion gestures to remotely manipulate 3D objects on public displays via a mobile device. Navas Medrano et al. [68] explored the use of mobile devices for remote deictic communication, and reported that most participants used either the (assumed) camera of the device or their fingers for pointing. Kray et al. [25] investigated the use of gestures to combine mobile phones with other devices (i.e., other mobile phones, interactive tabletops, and public displays). They indicated that users generally liked gestures to use their mobile phones to interact with other devices. Here also, one can see that mobile interaction with public displays is promising, and has already attracted interest from previous research. Nonetheless, similar to gestural interaction with public displays, previous work has not sufficiently addressed the design of intuitive mobile phone interaction for gathering citizens' feedback on large displays. Our work aims to fill this gap.

3. User Study

The main goal of our elicitation study is to investigate user-defined interactions to enable citizens to provide feedback about urban planning proposals on large public displays. We focused on four tasks: *vote for yes*, *vote for no*, *leave a comment*, and *delete a comment*. We were interested in finding answers to the following research questions:

3.1. Study Design

To provide a realistic setting for the elicitation study, we created two urban planning scenarios as shown in Figures 1 and 2 (The text in the speech balloon in Figure 2 is "The building is very nice, I like the color."). The reader is referred to Appendix A for a detailed description of the two urban planning scenarios. These scenarios each consisted of a single proposed building to avoid any confusion that could result from complex proposals. Our user study was performed in a single user setting in line with the scenario presented in Section 1. We thus simulated an asynchronous use of the public display by multiple users. Besides, it makes sense first to explore gestures preferred by individual users, before eliciting those relevant to the scenario of multiple concurrent uses. We presented the

two scenarios using an IVE, where we overlaid urban planning materials over the panoramic video footage. The IVE can immerse users in panoramic video footage and give them a feeling of “being ‘there’ rather than being ‘here’” [69]. This immersion can improve participants’ engagement, which we thought would further reduce barriers and facilitate gesture generation. We selected the two interaction modalities (hand gestures, smartphone gestures) based on the following considerations. First, we wanted to make sure that non-professionals could easily interact with the system. Therefore, we wanted to look for interaction methods that are intuitive and do not require technical expertise so that many citizens could potentially use them. Then, smartphones have become ubiquitous and thus might be a realistic and readily available means of interaction. Although multi-touch has also been considered to be an intuitive interaction modality [18], we did not include it in our study for three reasons. First, the public display in our study is very large and vertical so that reach-ability might have been a problem. Second, stepping up close to the display would have made it difficult for participants to perceive the entire scene they were voting/commenting on. Third, we wanted to avoid reducing the immersive user experience [70], which could negatively affect participation.

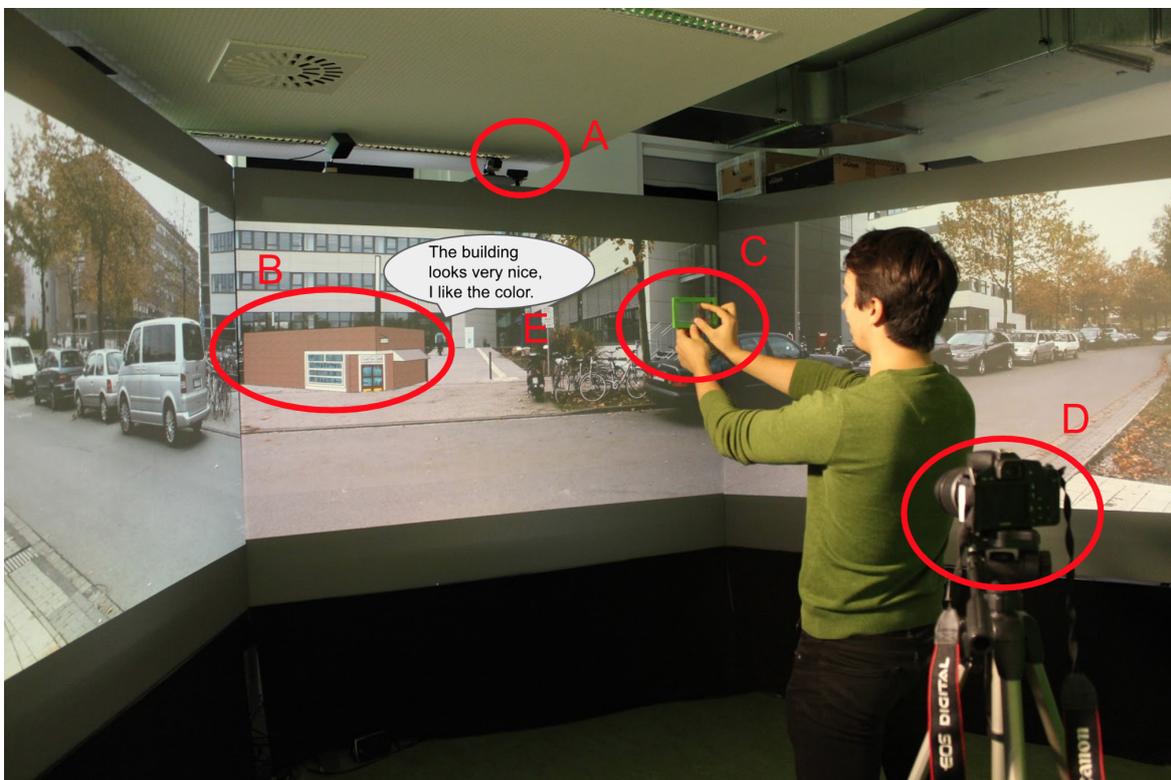


Figure 1. Study setup: a participant with the transparent mock-up prototype (C) standing on the footprint mark (not shown) in front of the IVE showing urban planning material (B) with an example comment (E), while a GoPro Hero4 camera (A) and a Canon camera (D) are recording their behavior.



Figure 2. An example stimulus used during our study.

In addition to the interaction modalities, we also had to select the tasks that participants had to perform. As previous work [17,45] revealed, in most cases, the primary function of current systems for public participation in urban planning is an informative one. Our research focuses on supporting the move from a lower level (i.e., informing) to a higher (i.e., consultation) on Arnstein's [26] ladder of citizen participation in urban planning. According to the Spectrum of Public Participation developed by the International Association of Public Participation, citizen consultation aims "to obtain public feedback on analysis, alternatives, decisions or both (<https://bit.ly/2lmPtTe>, last accessed: 6 February 2019.)". Two broad tasks are thus relevant in the context of public consultation: examining urban planning materials and giving feedback. In this paper, we are mainly focusing on the task of giving feedback; example gestures relevant to the examining task were synthesized in a previous paper [71]. Hanzl et al. [45] reviewed the forms of collecting citizen opinions in the urban planning context, and voting and commenting were found as two main forms of giving feedback.

Instead of providing participants with a specific phone model, we handed them a transparent (non-functional) mock-up prototype (as shown in Figure 1). We told the participants that the mock-up prototype was a "futuristic smartphone", which could have any features they could imagine. Our rationale for doing this mirrors the reasons mentioned in previous work [23,53,68]. First, we did not want the participants' behavior to be influenced by technical limitations such as gesture recognition issues or smartphone sensor technologies. Second, we want to minimize the effect of the gulf of execution [72] from the user-device dialogue. The study was approved by the institutional ethical review board.

3.2. Participants and Tasks

28 people participated in our study, of which seven were female. Participants were between 22 and 39 years old (mean = 28, SD = 4.9). We advertised our study through emails, Facebook, and word of mouth. Though the study took place in Germany, we conducted the study in English based on several considerations. The city where the study was run has many residents from diverse countries. Quite a few of them speak English at work (e.g., at the university or at one of the local companies) but do not speak German well. In addition, the English proficiency of the native population is very high. Therefore, picking English as the study language enabled us to include a broad range of users, including some that would otherwise have been excluded. English was also the language shared by all researchers who were involved in the study. At the end of the study, every participant received 10 Euros as compensation for their time. The whole study for each participant took around 40–60 min.

To keep our study focused and manageable, we defined four tasks that are essential for actively participating in an urban planning process. These tasks were: (1) leave a comment about the design of the building; (2) delete a comment from the large public display; (3) vote yes whether the building should be built as proposed; and (4) vote no whether the building should be built as proposed. All four tasks were executed for both modalities and were presented to the participants as a question, resulting in a total of eight questions per participant. Four questions stated: "Which hand gesture would you use to do ACTION?" and four questions asked: "How would you use your smartphone to do ACTION?". ACTION stands for one of the four tasks listed above.

3.3. Apparatus and Materials

The study took place in a lab environment, where an Immersive Video Environment (IVE) was installed. It consisted of three big screens, which were arranged in a semicircular way and which were connected to a single PC running on Windows 7. The PC played back the two stimuli (as shown in Figures 1 and 2) so that they spanned all three big screens. Each stimulus consisted of a panoramic video that showed a real-world location near our lab. This video was overlaid with a 3D building model, which was the proposed urban planning content. One building was a model of a supermarket and the other one the model of a skateboard shop. We exported the two 3D models from Sketchup Pro 2017 into PNG format with transparent backgrounds, which were then overlaid on the panoramic

videos using Final Cut Pro X. We also installed two video cameras to record all the details of the study. One camera was installed above the middle screen, recording the participant from the front, and one was mounted on a tripod to their left and slightly behind them (see Figure 1). During the study, every participant stood at the same location in the room, which was marked by one paper with two footprints. The mock-up smartphone consisted of two sheets of plexiglass that were attached by green adhesive tape. The tape ran around all the edges of the mock-up, leaving the central part free. Users could thus look through the mock-up. The size was similar to a regular smartphone.

3.4. Procedure

After welcoming a participant, the experimenter described the aims and the procedure of the study. Participants were then provided with a consent form they had to sign, while we also made sure that they understood the purpose of the study. Then participants were led to the IVE and asked to stand on the marked location (footsteps) so that they faced the three screens, with the middle one right in front of them. Once they were so positioned, the following steps were carried out:

1. The experimenter briefly described the urban planning story (as shown in the Appendix A.1) corresponding to the first stimulus.
2. Participants were given time to familiarize themselves with the stimulus.
3. Participants were allowed to ask any questions they might have, and once they were answered (or if they had none), the recorded audio instructions for the first task were played back.
4. Participants then performed the requested gesture, thinking aloud if they chose to do so.
5. After a gesture was performed, participants were asked to rate its easiness and appropriateness for this task.
6. The steps above were repeated for all four tasks.
7. Once the four tasks were completed for the first condition, the experimenter moved to the second stimulus and the second condition and repeated the procedure described above.
8. After completing the second condition, participants had to fill out questionnaires about their demographics, their general feedback and their attitudes regarding the interaction modalities.
9. At the end of the study, they could again ask questions, were paid for their participation and then discharged.

We used a within-subject study design and counterbalanced the order of exposure to the two conditions (hand gestures and phone gestures). For each condition, the order in which the tasks were given was randomized across participants. The order of the two videos/stimuli was also randomized across conditions and participants to avoid any order effects.

4. Results

The data collected includes videos and post-study questionnaires during which background information and further qualitative feedback on participants' experience were collected. We digitized all post-study paper questionnaires using SurveyMonkey, an online survey tool. In this way, we converted all paper information into electronic information for further analysis. Regarding the video data recorded during the main study session, we first screened them according to the taxonomies from Table 1. After that, detailed annotation and analysis of the data took place.

Table 1. Taxonomies used to analyze the user-defined gestures (drawn and extended from [23,24,53]).

Dimension	Sub-Category	Description
User-defined hand gestures		
Form	static	A static body posture is held after a preparation phase
	dynamic	The gesture contains the movement of one or more body parts during the stroke phase
Body parts	right hand	The gesture is performed with the right hand only
	left hand	The gesture is performed with the left hand only
	two hands	The gesture is performed with the two hands
Complexity	simple	The gesture consists of a single gesture
	compound	The gesture can be decomposed into several simple gestures
User-defined phone gestures		
Form	static pose	The hand pose is held in one position without any changes
	dynamic pose	The hand pose changes in one position
	static pose and path	The hand pose is held as hand moves
	dynamic pose and path	The hand pose changes as the hand moves
Body parts	one finger	The gesture is performed with one finger only
	two fingers	The gesture is performed with two fingers
	multi-fingers	The gesture is performed with more than two fingers
Complexity	simple	The gesture consists of a single gesture
	compound	The gesture can be decomposed into simple gestures
Spatial	See-through	Perform gestures while looking through the transparent screen of the device
	Device-pointing	Device used an extension of the user's arm to point at targets on the display
	Standard smartphone use	Device held in one hand, and one or two fingers used to simulate interaction

Out of 224 interactions events, we had to exclude three from further analysis due to participants not producing any hand gesture. Also, two interaction tasks were completed by touching the large display rather than by producing hand gestures. Given the low usage of touch-screen, we did not count these two interaction tasks in our analysis either. Finally, all user-defined interactions collected from our study involved three main categories of interactions: gestures (including hand gestures and phone gestures), voice, and both, i.e., the combination of voice and gestures, as shown in Table 2. We emphasized to every participant at the beginning of the study that: firstly they should only focus on the hand or smartphone interaction modality during the whole study and secondly their interactions would be supposed to take place in a public space where the environment could be unpredictable with varying factors such as noise level and audiences. Despite these instructions, some participants used voice interactions or gestures combined with voice while executing eight tasks. These voice interaction data from the participants were not included in the final analysis. There are two main phone gestures classified by input modalities: surface gesture and motion gesture. Only one participant performed two motion gestures for the tasks *vote for yes/no* while all the other user-defined phone gestures were surface gestures. Given the low usage of motion gestures, we only constructed the surface gesture taxonomy for the phone interaction modality.

Table 2. Distribution of use of interaction modalities.

Tasks Performed	Gesture	Voice	Both
All	178	9	32
Hand	83	0	24
Smartphone	95	9	8

4.1. Gesture Analysis Method

The user-defined hand gestures and phone gestures were analyzed according to our taxonomies, which was based on previous work about eliciting user-defined gesture interactions [23,24,53]. Though different taxonomies have been proposed in the literature, we only focused on dimensions that are most relevant for gestural interactions by hands or smartphone. We manually classified each

hand gesture along three dimensions: form, body parts, and complexity. Regarding phone gestures, we extended existing taxonomies with one more dimension (i.e., spatial) to cope with the richness of our data (see Table 1).

The *form dimension* in the hand-gesture taxonomy was adopted from Obaid [24] without modification. Gestures classified as static gestures imply that the gesture was kept in the same location without movement after a preparation phase and before the retraction phase [73]. In contrast, gestures classified as dynamic gestures imply that the gesture has one or more movements during the stroke phase (which is the phase between the preparation and retraction phase). The classification was applied separately to each hand in a two-hand gesture. In the phone gesture taxonomy, we included the sub-categories under the form dimension from surface gestures [23]. Compared with the form dimension in the taxonomy for user-defined hand gestures, the form dimension in the taxonomy for user-defined phone gestures was extended into four sub-categories including *static pose*, *dynamic pose*, *static pose and path*, and *dynamic pose and path*. We included these four sub-categories to capture more detailed information about surface gesture motions.

We modified the categories of the *body-parts dimension* also from [24] since our study involved only hands and no other body parts. We distinguished between right hand, left hand, and two hands. In the taxonomy of the phone gestures, we changed sub-categories of the body-parts dimension according to the characteristics of phone gestures, which described the number of fingers involved while performing gestures. The *complexity dimension* aims at classifying how complicated one gesture is. The compound gesture could be decomposed into simple gestures by segmenting around spatial discontinuities [53], (i.e., the corners, pauses in motion, or inflection points) or path discontinuities, (i.e., pauses or discontinuities in fingers' movement) in the gestures.

Form, body parts, and complexity were suggested as dimensions in previous work. To be able to cope with the richness of our data, we brought forth one more dimension, namely the *spatial dimension*. This dimension encodes how people hold or use the phone while interacting. We distinguish between three different ways: *see-through*, *device-pointing* and *standard smartphone use*. Some participants preferred performing gestures by holding the phone with two hands, and looking through the transparent screen of the mock-up device: these were classified as see-through (ST) gestures after [68]. Some participants also used the mock-up device as an extension of their arms to point at targets on the display, and this category was labeled device-pointing (DP). We also found that some participants designed gestures that mimicked actions occurring during standard smartphone use (i.e., the device was held in one of the hands, and one or two fingers were used to simulate interaction). These gestures were categorized as standard smartphone use (SSU). The main difference between ST and SSU is that participants systematically put the mobile phone in front of their eyes for ST, while they kept it in their hands regarding SSU.

4.2. Gesture Types and Frequency

Figure 3 shows the distribution of the 107 hand gestures according to their types. There were more static (60%) than dynamic (40%) left-hand gestures, while there were less static (47%) than dynamic (53%) right-hand gestures. However, the number of right-hand gestures (82) is larger than the two-hand gestures (10) or left-hand gestures (15). According to the information we collected from the background information of the participants, more than 90% were the right-handed person. This should be the reason there were more right-hand gestures than left-hand gestures. Most hand gestures tended to be simple gestures (66%). Figure 4 illustrates the distribution of all elicited phone gestures, i.e., surface gestures. Similar to the distribution of hand gestures, there were simpler (54%) than compound gestures (46%) though the difference was not so big. More than 90% of gestures were performed by one finger; however, the number of SSU (49%) gestures was almost equal to the number of ST gestures (50%).

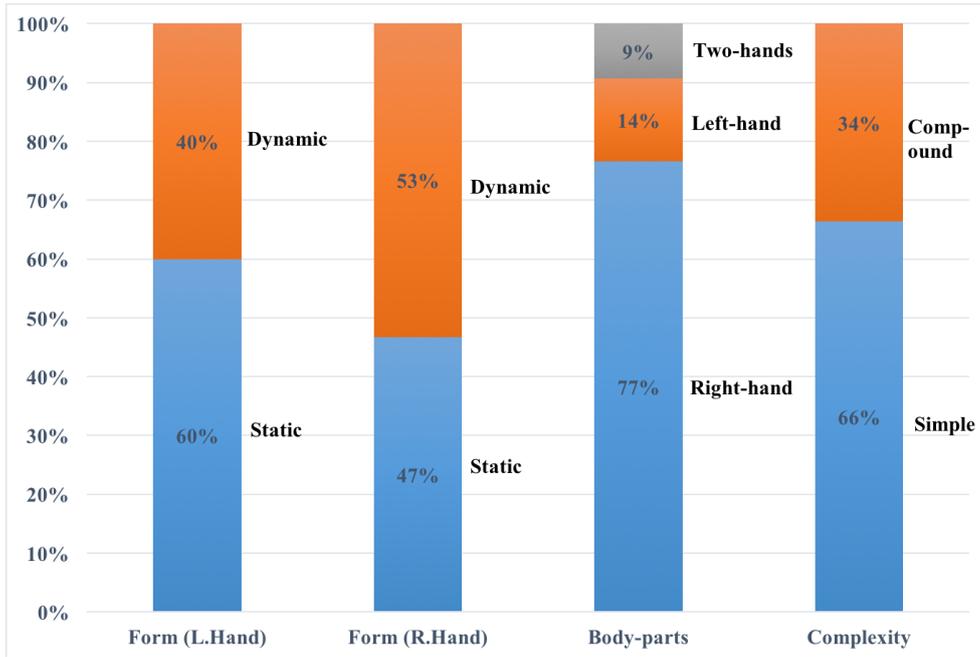


Figure 3. Distribution of the elicited hand gestures.

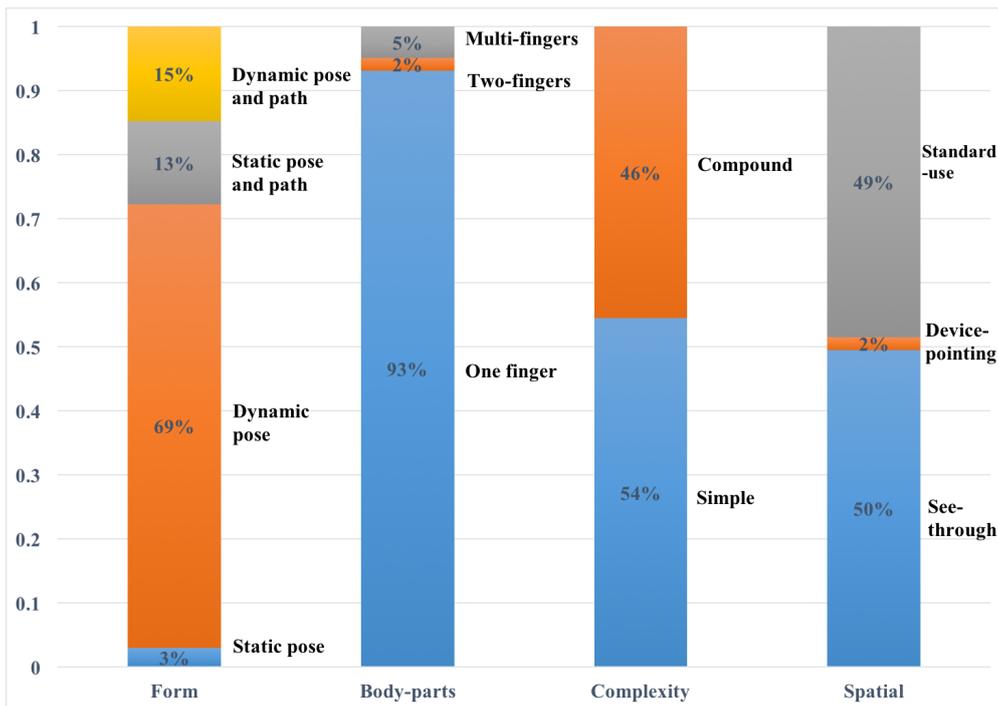


Figure 4. Distribution of the elicited phone gestures.

4.3. Gesture Sets

Based on the elicited gestures, another main result of our study consists of two user-defined gesture sets for the specified tasks by conditions: hand gestures and smartphone gestures. Firstly, we had the set of proposed gestures for each task t , which was known as $P(t)$. Secondly, for each task t , we grouped all identical gestures. We used $P_i(t)$ to represent a subset of identical gestures from $P(t)$. Thirdly, the subset $P_i(t)$ with the largest size was chosen as the representative gesture, i.e., gesture candidate, for the task t . To avoid the conflict of two tasks getting the same gesture candidate, we added the second gesture candidate, which was chosen only when it accounted for

least half of the first gesture candidate. Finally, the user-defined gesture set, which was also called the consensus set [53], was generated with the representative gestures of all the tasks. The resulting user-defined sets are shown in Tables 3 and 4.

Table 3. Gesture candidates for four tasks by hands.

Task	Gesture Candidates	Occurrence	Form	Body Parts	Complexity
Vote for Yes	thumb-up	64%	static	one hand	simple
Vote for No	thumb-down	61%	static	one hand	simple
Delete a comment	point to the comment, then swipe across the comment off-screen	36%	dynamic	one hand	compound
	draw a X sign	21%	dynamic	one hand	compound
Leave a comment	air-point to the display to trigger voice input	43%	static	one hand	simple

Table 4. Gesture candidates for four tasks by smartphone.

Task	Gesture Candidates	Occurrence	Form	Body Parts	Complexity	Spatial
Vote for Yes	tap the imaginary voting button	68%	dynamic pose	one finger	simple	SSU 58%, ST 37%, DP 5%
Vote for No	tap the imaginary voting button	68%	dynamic pose	one finger	simple	SSU 58%, ST 37%, DP 5%
Delete a comment	swipe across the comment off-screen	39%	static pose and path	one finger	simple	ST 55%, SSU 45%
Leave a comment	tap the imaginary comment button, then keyboard-type	46%	dynamic pose	multi-fingers	compound	ST 54%, SSU 46%

We also calculated the agreement rate (AR) [74] for each task to evaluate the degree of consensus among all the participants. Vatavu and Wobbrock [75] further advocated the use of this formula:

$$AR(t) = \frac{\sum_{i=1}^n P_i(t)(P_i(t) - 1)}{P(t)(P(t) - 1)}$$

Compared to the widely used agreement score introduced by [76], the agreement rates are more accurate measures of agreement [74,75]. And also for the within-participants design, the agreement rate is a valid approach [56]. The agreement rates for the two conditions are shown in Figure 5. The agreement rates for phone gestures are higher than the agreement rates for hand gestures, indicating higher consensus for tasks carried out by phone gestures.

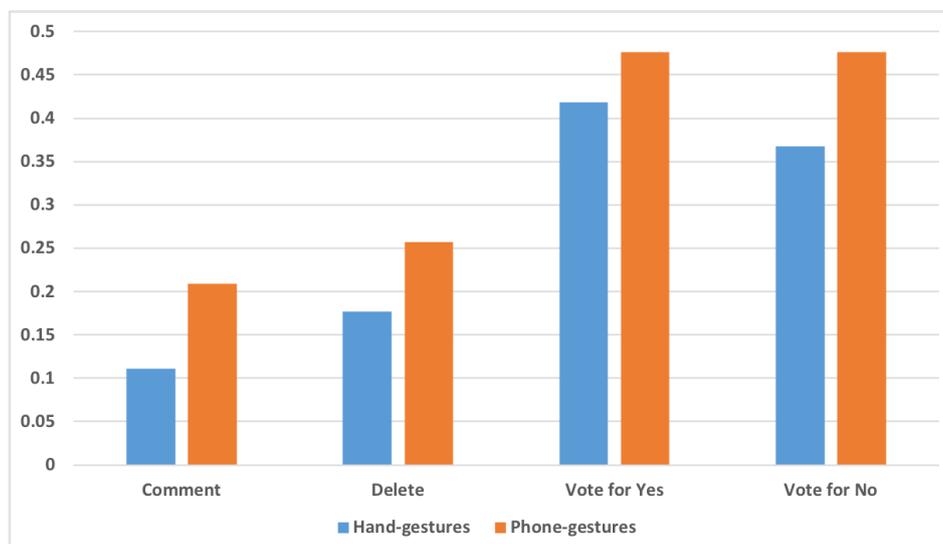


Figure 5. Agreement rates for the user-defined gestures.

4.4. User Ratings

We also asked participants to rate the appropriateness and easiness of their proposed gesture or action, after each task. Our goal was to assess the quality of the match of their gestures and tasks. Two questions were asked: (1) How easy was it for you to produce this gesture? (The scale given to participants ranged from 1 = quite hard to 5 = quite easy); and (2) How would you rate the appropriateness of your gesture/action to the task? (The scale given to participants ranged from 1 = quite inappropriate to 5 = quite appropriate). We used R to calculate the means of participants' ratings for these two questions. Table 5 shows the mean ratings of appropriateness and easiness in the two conditions. It is interesting to observe that the means for appropriateness and easiness of phone gestures for the tasks *leave a comment* and *delete a comment* are higher than those of hand gestures, while the means of appropriateness and easiness of phone gestures for the tasks *vote for yes* and *vote for no* are lower than those of hand gestures. All mean ratings of easiness for phones gestures for the four tasks are bigger than 4 (4 = easy; 5 = quite easy), while the mean ratings of easiness for hand gestures for the four tasks fall in the range of 3–5 (3 = moderate). Similarly, all mean ratings of appropriateness for phone gestures are around 4 (4 = appropriate; 5 = quite appropriate), but the mean ratings of easiness for hand gestures fall within the range of 3.5–4 (3 = moderate).

Also, a two-way repeated measures ANOVA revealed that the *ratings of easiness* did differ as a function of both tasks and interaction modalities, $F(3) = 4.5428$, $P < 0.01$. We also applied a two-way repeated measures ANOVA to the ratings of appropriateness, but no significance was found. By applying one-way repeated measures ANOVA, we found that the ratings of easiness differed significantly between the tasks with $F(3) = 6.978$, $P < 0.001$ for hand interactions, but not for phone interactions. In particular, the task *leave a comment* received significantly lower ratings than the task *vote for yes* with $P < 0.001$ and *vote for no* with $P < 0.01$ for the hand interaction modality. No significant difference was found for *delete a comment*. Similarly, no significant difference was found among the tasks regarding the *ratings of appropriateness*, either for hand or phone interactions.

Table 5. Appropriateness and Easiness ratings for the gestures proposed (ME, means of easiness; SDE, standard deviation of easiness; MA, means of appropriateness; SDA, standard deviation of appropriateness).

Task	Hand Gesture				Phone Gesture			
	ME	SDE	MA	SDA	ME	SDE	MA	SDA
Comment	3.25	1.37	3.46	1.01	4.21	0.98	4.07	0.65
Vote for Yes	4.39	1.00	4.14	1.09	4.04	1.12	4	0.89
Vote for No	4.39	1.00	4.14	1.09	4.04	1.12	4	0.89
Delete	3.79	0.85	3.79	0.85	4.04	1.12	3.89	0.77

4.5. User Feedback

At the end of the study, we asked for positive and negative aspects of using hands and phones as interaction modalities for giving feedback on the display. We read and re-read all text answers and looked for common themes among them. Using Excel, we then counted the number of responses that applied to each theme. In the following, we list the most frequent positive/negative feedback themes (a “most frequent feedback” was mentioned by at least two users).

The four most frequent positive comments about *hand gestures* were: no need for a personal device; more intuitive and natural; more fun and attractive to use; more accessible for broader groups. Participants pointed out for example that hand gestures are “not needing an additional device”, that there is “no need to bring phone out from pocket”; that they are “more natural”, and “more fun, due to their game-play-effect”; also that they “can be more intuitive” and that “anyone will be able to join in”. Regarding the *phone interaction modality*, the three most frequent positive comments were: more familiar with phone interactions; faster and easier; more private and less social

embarrassment. Here, participants indicated, for instance, that interaction using phone gestures is “easy, fast, and familiar”; that “the generating of interactions are easier because people are taught by certain apps and smartphone interaction how to use interactions like making things bigger or smaller”. They also pointed out that “it is more private” and that “people have to perform less ‘embarrassing’ gestures in front of people. Also, the interaction can be made more accurate and potentially faster. Multiple users could add comments”.

The four most frequent negative comments regarding *hand gestures* included being unfamiliar with hand gestures; insensitive and inaccurate gesture recognition; the social embarrassment of performing gestures in public; no privacy. Here, participants stated: “not easy to know how to use. Gesture detection can be complicated to detect. It can also be harder to define different gestures for similar actions. Frustration from the user if the system doesn’t detect gestures properly”; that it “does not feel familiar, might feel strange to move like this in public”; that it comes with a “lack of privacy, everyone can see me perform, possible to clash with other people performing gestures”. The four most frequent negative comments for the *phone interaction modality* cover the need for an additional personal device; more effort involved in switching between two screens; extra effort to connect a personal device with the display; and the limited screen size of the smartphone. Participants perceived the “use of the extra personal device. Might require to download a special app to be able to interact”; “more effort to check two screens (large and small)”; “the device needs some connection to the public display” and “small screen size is not very comfortable”; as negative.

At the end of the study, each participant was also asked the following question: “Imagine a system like the one you used would be deployed in the real world to gather feedback on proposed building projects. Would you want to comment on such projects using large displays installed in public locations? Why? or Why not?” (cf. Appendix B.2). More than 90% of participants confirmed their intention to use such a system if it was deployed.

5. Discussion

Overall, feedback from participants for using smartphone or hand gestures for giving feedback on urban planning materials was positive. However, according to Figure 5, the agreement rates for user-defined phone gestures are higher than the agreement rates for user-defined hand gestures. We also observed that participants’ preference for two interaction modalities were substantially affected by the type of tasks. Participants could represent actions with hand interactions more easily and quickly. They also thought the hand interaction proposed was more appropriate regarding the tasks *vote for yes* and *vote for no*. However, they performed better with the smartphone interaction and thought the smartphone interaction proposed was more appropriate regarding the tasks *leave a comment* and *delete a comment*.

Revisiting the three research questions listed in Section 3, the user-defined sets which are reported in Tables 3 and 4 provide an answer to the first research question. Providing an answer to the second research question (“*which of these four tasks lend themselves well to being executed by hand or phone interaction modalities, and which ones do not?*”) depends on the definition of what “lending itself well” means. We adopted the notion that tasks which ‘lend themselves well’ for an interaction modality need at least a score of 4.0 for that interaction modality. This appeared to be the best match to the question asked to the participants and the Likert scale given to them (as a reminder 3.0 meant ‘moderate’, 4.0 meant ‘appropriate’, and 5.0 meant ‘quite appropriate’ on that scale). Looking at Table 5, one can see that the tasks *leave a comment* and *delete a comment* lend themselves well to being executed by the phone interaction modality in our study (the hand interaction modality did not reach the threshold set at the beginning of the paragraph).

In contrast to the above, we obtained quite different results for the tasks *vote for yes* and *vote for no*. From the two figures, the mean ratings of appropriateness and easiness for the tasks *vote for yes* and *vote for no* are higher than 4, while the means for phone gestures are around 4, i.e., a little higher or lower than 4. This indicates that the tasks *vote for yes* and *vote for no* lend themselves well to being

executed by both of the two interaction modalities (though the hand interaction is likely to be users' first choice in practice since it received overall higher ratings during the study).

Both interaction modalities have their advantages and disadvantages on specific tasks as discussed in the previous paragraph. However, regarding the answer to the third research question, we can state that the phone interaction modality may be more suitable than the hand interaction modality for the four tasks, based on the results of this study. The primary rationale is that the mean ratings of appropriateness and easiness for phone gestures always kept around 4, which suggests that although the phone interaction modality did not perform as good as the hand interaction modality on the tasks *vote for yes* and *vote for no*, its overall performance was better. In addition, for the four tasks, the phone interaction was not only considered to be easy to come up with but also appropriate for the tasks. Regarding the hand interaction modality, it executed the tasks *vote for yes* and *vote for no* better than the phone interaction modality based on the results of easiness and appropriateness ratings, but the difference was minimal. For the task *leave a comment*, the mean rating of appropriateness was around 3.5 and the mean rating of easiness was between 3.5 and 3, which shows a big difference for the phone interaction modality. The mean ratings of appropriateness and easiness of hand gestures for task *delete a comment* were lower than 4.

5.1. Implications for System Design

The user-defined gesture sets listed in Tables 3 and 4 and some trends that emerged during the study provide several challenges for designers of gestural interaction with the public display. One trend that emerged is that not only gesture recognition sensors, but also sound sensors may be needed to collect feedback from participants on public displays in an urban planning context. 20/28 participants used voice or voice+gestures as input to leave their comments when using hand gestures, which means sound sensors should be sensitive enough to capture people's words. At the beginning of the study, every participant was told that she/he was expected to perform all the tasks/actions through hand gestures or smartphones. However, many ended up using voice along with their proposed gestures. There may be at least three reasons for that: (i) it may indicate that people prefer to use hand gestures to perform direct and simple actions, e.g., *vote for yes* and *vote for no*, while they would like to use voice as an add-on to hand gestures for the complex tasks, e.g., *leave a comment* (see Table 3); (ii) it may be an indication that interaction with large display should naturally come up with voice as input modality; (iii) it may be a consequence of legacy bias (discussed briefly in Section 5.6). Testing (i), (ii) and (iii) necessitates further investigations in future work. Another major trend which emerged was that participants performed gestures by looking through the transparent screen of the mock-up device. This may indicate that people do not want to make complicated connections between the smartphone and the display. Instead, they mimic the connecting process as a way similar to the process of scanning the QR code. In our study, we found that people expected real-time, synchronous responses to their interactions on the smartphone and the large display.

Furthermore, participants' comments on the drawbacks of the interaction modalities should be taken into account by interface designers to minimize risks of interaction blindness [77,78] in the urban planning context. That "lack of privacy" was mentioned by several participants as a possible negative aspect, points at the need for further researching methods (such as [79]) to guarantee a "personal space" to users while interacting with public displays for voting/commenting in the urban planning context. Additionally, the immersive display explored in our study arguably presents more restricted access to the displayed contents than other large displays deployed in completely public spaces, e.g., the displays in a plaza or on a street. However, this kind of feature is consistent with citizens' potential privacy concerns mentioned in their comments. The privacy concerns could be mitigated by deploying immersive displays to gather their feedback in semi-public spaces (e.g., a controlled environment such as an office building of the local city council). This kind of semi-public space also provides the possibility for 'personal space' concerning social and privacy impacts [80], especially in the context of citizen voting and commenting. In this case, local authorities could provide a smartphone in these

spaces as a way to reduce potential hesitations of some citizens to use their personal devices. However, using personal devices to interact with the immersive display is also an option. According to the work of Baldauf et al. [81], one alternative approach to mitigate privacy concerns could be to combine personal devices with public displays. Personal information could be overlaid on users' own devices via AR while they are being pointed at a public display.

5.2. Implications for ICT-Based Citizen Participation

Table 5 illustrates that the participants overall found it easy to produce the gestures for the four tasks considered. Since gesture-elicitation aims primarily at finding out a vocabulary for interaction with devices, these results suggest that voting/commenting on public displays using gestures could be easy (and thus worth implementation efforts into prototypes) in the citizen participation context. It should be noted that whether these tasks end up being as easy as suggested by the results require full prototype developed and testing in subsequent studies. For example, in another study [82], participants have been asked to use a simple prototypical mobile client on a smartphone to vote/comment on urban planning contents, and reported high values for usefulness, ease of use, ease of learning and user satisfaction. The design of the prototype was informed by the results of this study. In sum, Table 5 provides evidence that ICT-based citizen participation via public displays indeed holds potential in the digital age.

5.3. Gesture Consistency and Reuse

We observed that participants occasionally reused simple gestures for different tasks. For example, the gesture candidates for the task *vote for yes* and *vote for no* by hands were thumb-up and thumb-down, respectively, while the gesture candidates for these two tasks by smartphone were similar (i.e., tap the imaginary voting buttons). Participants reused some of their gestures not only among analogous tasks in the same condition but among different conditions for the same task. For example, more than half of participants used one hand to swipe across the comment off-screen from the display, and more than half participants chose to swipe across the comment off-screen from the smartphone by one finger. This emphasizes people's natural propensity to reduce learning and memorizing workload. Gesture reuse and consistency are thus important to increase learnability and memorability [23,60,83], and this should be taken into account when designing user interfaces to gather citizens feedback in an urban planning context.

5.4. Interaction beyond Single Interaction Modality

Figure 5 shows the agreement for the task *leave a comment* by hands is low, which suggests that participants found it challenging to complete the task with a gesture or an interaction. As said above 20/28 participants chose voice or voice+hand gestures to leave their comments instead of hand gestures only. When they performed the tasks with the smartphone, some of them (8/28) also used voice or voice+phone gestures to finish the task of leaving their comments. Few people even insisted on only using voice interaction with the smartphone. This indicates that voting/commenting in the urban planning context may benefit from a combination of the two interaction modalities.

5.5. Beyond Public Consultation in Urban Planning

Though the scenario used during this work has been the public consultation in urban planning, voting, and commenting are also important in other contexts. In particular, two application scenarios of public display mentioned in [18] could be equally relevant, namely learning and gaming. For instance, students could be asked to give feedback (like/dislike or comments) on a learning material in an informal learning context [84], and gamers could be asked to rate their gaming experience (like/dislike) before leaving an interaction session. The hand gestures and smartphone-based gestures collected in this study can be useful in both scenarios (and virtually in any context where collecting feedback on public displays can be reduced to voting/commenting tasks similar to this work).

5.6. Limitations

Overall, we found participants to be very creative to generate many hand gestures in a very limited time. According to the background information on participants collected, almost 90% of them had rarely interacted with a public display by using hand gestures. However, because of the limitation of the elicitation methodology, there may be legacy bias [49] in our study, which means that the user-defined smartphone interactions were quite influenced by their experience of using smartphones during daily life. Around 96% of participants declared that they used smartphones very frequently in their daily life. This is also reflected in their generation of the user-defined set for smartphone interaction. So even though users participated freely voiced their preferences, it may be the case that the interactions suggested have been slightly biased by (or towards) their smartphone using experience. Including more participants with less familiarity with smartphones in subsequent studies could help to get a complete picture of gestural interactions for voting/commenting. Our study considered a single user setting (i.e. asynchronous use), but eliciting user gestures for concurrent use by multiple users appears could be another promising extension to investigate not only *how* multiple users could interact with the display, but also *how many* can simultaneously interact. Finally, the fact that some participants suggested voice (alone or in combination with gestures) as an interaction modality during the study—without being asked to do so—is an indication that the combined use of gestures+voice for public display interaction in urban planning scenarios needs a closer look in future studies.

6. Conclusions

A total of 219 interaction data from 28 participants were collected during an elicitation study on voting/commenting urban planning material on public displays. Participants came up mostly with one-hand gestures to interact with the public display using their hands (e.g., thumb-up, thumb-down as hand gestures to vote by hands, pointing to the comment and then swiping across the comment off-screen to delete it); they produced one-finger and multi-finger gestures regarding public display interaction with the smartphone (e.g., tap an imaginary button and then keyboard-type to delete a comment, or tap an imaginary voting button to vote “Yes”). Finally, our data suggest that smartphones would be more suitable for adding/deleting comments, while hand gestures would be more convenient for the voting task.

Implementing the smartphone gestures identified and evaluating their usability in the context of urban planning is part of ongoing work (see e.g., [82]). The previous experience of participants with the use of smartphones could have influenced the gestures they proposed, and a complementary elicitation study involving people with less familiarity with smartphones would be valuable in confirming/expanding the gesture set identified. Finally, since the work focused on the interaction with the public display for a single user, another possible extension of the work would be to conduct an elicitation study to collect user gestures for concurrent use (e.g., collaborative addition/removal of comments on the urban planning materials) of the public display.

Author Contributions: G.D. designed and conducted a user study, data analysis, and drafted most part of the manuscript. A.D. and C.K. guided the study, reviewed the manuscript, and assisted with the manuscript writing.

Funding: This research was funded by the European Union through the GEO-C project (H2020-MSCA-ITN2014, <http://www.geo-c.eu/>) under Grant number 642332.

Acknowledgments: We thank the anonymous reviewers for their comments, which have been helpful in improving the article.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ICT	Information and communication technologies
AR	Augmented reality
VR	Virtual reality
ME	Means of easiness
SDE	Standard deviation of easiness
MA	Means of appropriateness
SDA	Standard deviation of appropriateness
ST	See-through
DP	Device-pointing
SSU	Standard smartphone use

Appendix A. Urban Planning Scenarios Used during the User Study

By reviewing urban planning projects of the authors' city over the years and analyzing the current status of the authors' working place and the city, we proposed two urban planning scenarios. The two scenarios described below with the considerations: firstly we want to use the scenarios, which people are familiar with and often occur in the modern cities; secondly, we want to simulate different levels of urban planning projects: the community level project (i.e., scenario Appendix A.1), and the city level project (i.e., scenario Appendix A.2).

Appendix A.1. Scenario 1—A Skateboard Shop in Your University

Young people wish to have a skateboard shop in the campus, therefore they can buy skateboards easily without going to the city center, where is far away from the city center. There is the initial design plan for the shop. As a student, your participation of the design of the skateboard shop is very important and necessary. Your feedback will influence the final decision on its design. Your feedback can be given in two ways: voting and commenting.

Appendix A.2. Scenario 2—A Supermarket Next to the Train Station

Our city urgently needs a supermarket next to the train station, where people can easily buy food and drinks before or after a traveling. So the whole city decides to build a new supermarket. There is the initial design plan for the supermarket. As a citizen, your participation of the design of the supermarket is very important and necessary. Your feedback will influence the final decision on its design. Your feedback can be given in two ways: voting and commenting.

Appendix B. Questionnaires Used in the Study

There were two types of questionnaires in our study: the questionnaires for the participants' background information and the questionnaires for their attitude regarding the interaction modalities. Please find them as follows.

Appendix B.1. The Background Information Questionnaire



Questionnaire

ID _____

1. How often do you use a smartphone?

- Never Rarely Occasionally Frequently Very frequently

2. What are the operation systems of your devices (both smartphone and computer) you are using now?

- Windows MacOS/iOS Android Linux/Unix other _____

3. How often do you pay attention with a public display?

- Never Rarely Occasionally Frequently Very frequently

4. Have you interacted with a public display by using hand gestures?

- Never Rarely Occasionally Frequently Very frequently

5. How familiar are you with the area of the main train station of Muenster?

- Not at all Barely Moderate Well Quite well

6. How familiar are you with the area of GEO 1, heisenbergstraße 2?

- Not at all Barely Moderate Well Quite well

7. How often do you participate public events about urban planning in last 2 years?

- Not at all Barely Moderate Well Quite well

8. Your gender?

- Female Male

9. What kind of person are you, right-handed or left handed? _____

10. How old are you? _____

11. What is your profession? _____

12. What has been your main country of residence for the past two years? _____

We would like to thank you for your participation!

Appendix B.2. The Open Questions for Participants' Attitude



Open questions

ID _____

- 1. Of the two interaction methods you used (hand gestures, smartphone), which one do you like better?**

- 2. What negative aspects do you see when using hand gestures to interact with urban planning Information?**

- 3. What positive aspects do you see when using hand gestures to interact with urban planning Information?**

- 4. What negative aspects do you see when using the smartphone to interact with urban planning Information?**

- 5. What positive aspects do you see when using the smartphone to interact with urban planning Information?**

- 6. Imagine a system like the ones you used would be deployed in the real world to gather feedback on proposed building projects. Would you want to comment on such projects using large displays installed in public locations? Why? or Why not?**

References

1. Kaviani, N.; Finke, M.; Fels, S.; Lea, R.; Wang, H. What goes where?: Designing interactive large public display applications for mobile device interaction. In Proceedings of the First international Conference on internet Multimedia Computing and Service, Kunming, China, 23–25 November 2009; ACM: New York, NY, USA, 2009; pp. 129–138.
2. Müller, J.; Walter, R.; Bailly, G.; Nischt, M.; Alt, F. Looking glass: A field study on noticing interactivity of a shop window. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, TX, USA, 5–10 May 2012; ACM: New York, NY, USA, 2012; pp. 297–306.
3. Peltonen, P.; Kurvinen, E.; Salovaara, A.; Jacucci, G.; Ilmonen, T.; Evans, J.; Oulasvirta, A.; Saarikko, P. It's Mine, Don't Touch!: Interactions at a large multi-touch display in a city centre. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy, 5–10 April 2008; ACM: New York, NY, USA, 2008; pp. 1285–1294.
4. Valkanova, N.; Jorda, S.; Tomitsch, M.; Vande Moere, A. Reveal-it!: The impact of a social visualization projection on public awareness and discourse. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 27 April–2 May 2013; ACM: New York, NY, USA, 2013; pp. 3461–3470.
5. Kuikkaniemi, K.; Jacucci, G.; Turpeinen, M.; Hoggan, E.; Müller, J. From space to stage: How interactive screens will change urban life. *Computer* **2011**, *44*, 40–47. [[CrossRef](#)]
6. Hinrichs, U.; Carpendale, S.; Valkanova, N.; Kuikkaniemi, K.; Jacucci, G.; Moere, A.V. Interactive public displays. *IEEE Comput. Graph. Appl.* **2013**, *33*, 25–27. [[CrossRef](#)] [[PubMed](#)]
7. Howard, T.L.; Gaborit, N. Using virtual environment technology to improve public participation in urban planning process. *J. Urban Plan. Dev.* **2007**, *133*, 233–241. [[CrossRef](#)]
8. Allmendinger, P.; Prior, A.; Raemaekers, J. *Introduction to Planning Practice*; John Wiley & Sons Inc.: Hoboken, NJ, USA, 2000.
9. Laurini, R. *Information Systems for Urban Planning: A Hypermedia Cooperative Approach*; CRC Press: Boca Raton, FL, USA, 2014.
10. El Araby, M. Possibilities and constraints of using virtual reality in urban design. In Proceedings of the 7th International CORP Symposium, Vienna, Austria, 27 February–1 March 2002; pp. 457–463.
11. Hosio, S.; Goncalves, J.; Kostakos, V.; Riekkki, J. Exploring civic engagement on public displays. In *User-Centric Technology Design for Nonprofit and Civic Engagements*; Saeed, S., Ed.; Springer International Publishing: Basel, Switzerland, 2014; pp. 91–111.
12. Ostkamp, M.; Kray, C. Supporting design, prototyping, and evaluation of public display systems. In *Proceedings of the 2014 ACM SIGCHI Symposium on Engineering Interactive Computing Systems—EICS '14*; Paternò, F., Santoro, C., Ziegler, J., Eds.; ACM Press: Rome, Italy, 2014; pp. 263–272.
13. Prouzeau, A.; Bezerianos, A.; Chapuis, O. Visual Immersion in the Context of Wall Displays. In Proceedings of the Interactive Surfaces and Spaces Surfaces Companion, Niagara Falls, ON, Canada, 6–9 November 2016.
14. Dede, C. Immersive interfaces for engagement and learning. *Science* **2009**, *323*, 66–69. [[CrossRef](#)]
15. Barab, S.A.; Sadler, T.D.; Heiselt, C.; Hickey, D.; Zuiker, S. Relating narrative, inquiry, and inscriptions: Supporting consequential play. *J. Sci. Educ. Technol.* **2007**, *16*, 59–82. [[CrossRef](#)]
16. Neulight, N.; Kafai, Y.B.; Kao, L.; Foley, B.; Galas, C. Children's participation in a virtual epidemic in the science classroom: Making connections to natural infectious diseases. *J. Sci. Educ. Technol.* **2007**, *16*, 47. [[CrossRef](#)]
17. Du, G.; Degbelo, A.; Kray, C. Public displays for public participation in urban settings: A survey. In Proceedings of the 6th ACM International Symposium on Pervasive Displays, Lugano, Switzerland, 7–9 June 2017; ACM: New York, NY, USA, 2017; pp. 17–25.
18. Ardito, C.; Buono, P.; Costabile, M.F.; Desolda, G. Interaction with large displays: A survey. *ACM Comput. Surv. (CSUR)* **2015**, *47*, 46. [[CrossRef](#)]
19. Nancel, M.; Wagner, J.; Pietriga, E.; Chapuis, O.; Mackay, W. Mid-air pan-and-zoom on wall-sized displays. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems—CHI '11*; ACM Press: Vancouver, BC, Canada, 2011; pp. 177–186.

20. Ackad, C.; Clayphan, A.; Tomitsch, M.; Kay, J. An in-the-wild study of learning mid-air gestures to browse hierarchical information at a large interactive public display. In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Osaka, Japan, 7–11 September 2015; ACM: New York, NY, USA, 2015; pp. 1227–1238.
21. Vanoni, D. Human-Centered Interfaces for Large, High-Resolution Visualization Systems. *Res. Exam* **2013**.
22. Poushter, J.; Bishop, C.; Chwe, H. *Social Media Use Continues to Rise in Developing Countries, but Plateaus Across Developed Ones*; Technical Report; Pew Research Center: Washington, DC, USA, 2018.
23. Wobbrock, J.O.; Morris, M.R.; Wilson, A.D. User-defined gestures for surface computing. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 4–9 April 2009; ACM: New York, NY, USA, 2009; pp. 1083–1092.
24. Obaid, M.; Kistler, F.; Kasparavičiūtė, G.; Yantaç, A.E.; Fjeld, M. How would you gesture navigate a drone?: A user-centered approach to control a drone. In Proceedings of the 20th International Academic Mindtrek Conference, New York, NY, USA, 17–18 October 2016; ACM: New York, NY, USA, 2016; pp. 113–121.
25. Kray, C.; Nesbitt, D.; Dawson, J.; Rohs, M. User-defined gestures for connecting mobile phones, public displays, and tabletops. In Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services, Lisbon, Portugal, 7–10 September 2010; ACM: New York, NY, USA, 2010; pp. 239–248.
26. Arnstein, S. A ladder of citizen participation. *J. Am. Inst. Plan.* **1969**, *35*, 216–224. [[CrossRef](#)]
27. Rowe, G.; Frewer, L.J. A typology of public engagement mechanisms. *Sci. Technol. Hum. Values* **2005**, *30*, 251–290. [[CrossRef](#)]
28. Godschalk, D.; Parham, D.; Porter, D.; Potapchuk, W.; Schukraft, S. *Pulling Together: A Planning and Development Consensus-Building Manual*; Urban Land Institute: Washington, DC, USA, 1994.
29. Innes, J.E. Planning through consensus building: A new view of the comprehensive planning ideal. *J. Am. Plan. Assoc.* **1996**, *62*, 460–472. [[CrossRef](#)]
30. Forester, J. *The Deliberative Practitioner: Encouraging Participatory Planning Processes*; MIT Press: Cambridge, MA, USA, 1999.
31. Straus, D. Designing a consensus building process using a graphic road map. In *The Consensus Building Handbook*; Sage Publications: Thousand Oaks, CA, USA, 1999; pp. 137–168.
32. Wondolleck, J.M.; Yaffee, S.L. *Making Collaboration Work: Lessons from Innovation in Natural Resource Management*; Island Press: Washington, DC, USA, 2000.
33. Lowry, K.; Adler, P.; Milner, N. Participating the public: Group process, politics, and planning. *J. Plan. Educ. Res.* **1997**, *16*, 177–187. [[CrossRef](#)]
34. Innes, J.E.; Booher, D.E. Reframing public participation: Strategies for the 21st century. *Plan. Theory Pract.* **2004**, *5*, 419–436. [[CrossRef](#)]
35. Gudowsky, N.; Bechtold, U. The role of information in public participation. *J. Public Delib.* **2013**, *9*, 3.
36. Sanoff, H. *Community Participation Methods in Design and Planning*; John Wiley & Sons: Hoboken, NJ, USA, 2000.
37. Creighton, J.L. *The Public Participation Handbook: Making Better Decisions through Citizen Involvement*; John Wiley & Sons: Hoboken, NJ, USA, 2005.
38. Krek, A. Rational ignorance of the citizens in public participatory planning. In Proceedings of the 10th Symposium on Information-and Communication Technologies (ICT) in Urban Planning and Spatial Development and Impacts of ICT on Physical Space, CORP, Vienna, Austria, 22–25 February 2005; Volume 5, p. 420.
39. Veenstra, M.; Wouters, N.; Kanis, M.; Brandenburg, S.; te Raa, K.; Wigger, B.; Moere, A.V. Should public displays be interactive? Evaluating the impact of interactivity on audience engagement. In Proceedings of the 4th International Symposium on Pervasive Displays—PerDis '15, Saarbrücken, Germany, 10–12 June 2015; Gehring, S., Krüger, A., Alt, F., Taylor, N., Eds.; ACM Press: Saarbrücken, Germany, 2015; pp. 15–21. [[CrossRef](#)]
40. Steinberger, F.; Foth, M.; Alt, F. Vote with your feet: Local community polling on urban screens. In Proceedings of the International Symposium on Pervasive Displays, Copenhagen, Denmark, 5–6 June 2014; ACM: New York, NY, USA, 2014; p. 44.
41. Hespanhol, L.; Tomitsch, M.; McArthur, I.; Fredericks, J.; Schroeter, R.; Foth, M. Situated interfaces for engaging citizens on the go. *Interactions* **2016**, *23*, 40–45. [[CrossRef](#)]

42. Wouters, N.; Huyghe, J.; Vande Moere, A. OpenWindow: Citizen-controlled content on public displays. In Proceedings of the 2nd ACM International Symposium on Pervasive Displays—PerDis '13, Mountain View, CA, USA, 4–5 June 2013; ACM Press: Mountain View, CA, USA, 2013; pp. 121–126.
43. Goncalves, J.; Hosio, S.; Liu, Y.; Kostakos, V. Eliciting situated feedback: A comparison of paper, web forms and public displays. *Displays* **2014**, *35*, 27–37. [[CrossRef](#)]
44. Mahyar, N.; Burke, K.J.; Xiang, J.E.; Meng, S.C.; Booth, K.S.; Girling, C.L.; Kellett, R.W. UD Co-Spaces: A Table-Centred Multi-Display Environment for Public Engagement in Urban Design Charrettes. In Proceedings of the 2016 ACM on Interactive Surfaces and Spaces, Niagara Falls, ON, Canada, 6–9 November 2016; ACM: New York, NY, USA, 2016; pp. 109–118.
45. Hanzl, M. Information technology as a tool for public participation in urban planning: A review of experiments and potentials. *Des. Stud.* **2007**, *28*, 289–307. [[CrossRef](#)]
46. Dambruch, J.; Krämer, M. Leveraging public participation in urban planning with 3D web technology. In Proceedings of the 19th International ACM Conference on 3D Web Technologies, Vancouver, BC, Canada, 8–10 August 2014; ACM: New York, NY, USA, 2014; pp. 117–124.
47. van Leeuwen, J.P.; Hermans, K.; Jylhä, A.; Qanjer, A.J.; Nijman, H. Effectiveness of Virtual Reality in Participatory Urban Planning: A Case Study. In Proceedings of the 4th Media Architecture Biennale Conference, Beijing, China, 13–16 November 2018; ACM: New York, NY, USA, 2018; pp. 128–136.
48. Hayek, U.W. Which is the appropriate 3D visualization type for participatory landscape planning workshops? A portfolio of their effectiveness. *Environ. Plan. B Plan. Des.* **2011**, *38*, 921–939. [[CrossRef](#)]
49. Morris, M.R.; Danieleescu, A.; Drucker, S.; Fisher, D.; Lee, B.; Wobbrock, J.O. Reducing legacy bias in gesture elicitation studies. *Interactions* **2014**, *21*, 40–45. [[CrossRef](#)]
50. Schuler, D.; Namioka, A. *Participatory Design: Principles and Practices*; CRC Press: Boca Raton, FL, USA, 1993.
51. Vatavu, R.D. User-defined gestures for free-hand TV control. In Proceedings of the 10th European Conference on Interactive TV and Video, Berlin, Germany, 4–6 July 2012; ACM: New York, NY, USA, 2012; pp. 45–48.
52. Piumsomboon, T.; Clark, A.; Billinghamurst, M.; Cockburn, A. User-defined gestures for augmented reality. In *Proceedings of the IFIP Conference on Human-Computer Interaction*; Springer: Basel, Switzerland, 2013; pp. 282–299.
53. Ruiz, J.; Li, Y.; Lank, E. User-defined motion gestures for mobile interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada, 7–12 May 2011; ACM: New York, NY, USA, 2011; pp. 197–206.
54. Chan, E.; Seyed, T.; Stuerzlinger, W.; Yang, X.D.; Maurer, F. User elicitation on single-hand microgestures. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; ACM: New York, NY, USA, 2016; pp. 3403–3414.
55. Hoff, L.; Hornecker, E.; Bertel, S. Modifying gesture elicitation: Do kinaesthetic priming and increased production reduce legacy bias? In Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction, Eindhoven, The Netherlands, 14–17 February 2016; ACM: New York, NY, USA, 2016; pp. 86–91.
56. Tsandilas, T. Fallacies of agreement: A critical review of consensus assessment methods for gesture elicitation. *ACM Trans. Comput.-Hum. Interact. (TOCHI)* **2018**, *25*, 18. [[CrossRef](#)]
57. Köpsel, A.; Bubalo, N. Benefiting from legacy bias. *Interactions* **2015**, *22*, 44–47. [[CrossRef](#)]
58. Groenewald, C.; Anslow, C.; Islam, J.; Rooney, C.; Passmore, P.; Wong, W. Understanding 3D mid-air hand gestures with interactive surfaces and displays: A systematic literature review. In *Proceedings of the 30th International BCS Human Computer Interaction Conference: Fusion!, HCI '16*; BCS Learning & Development Ltd.: Swindon, Poole, UK, 2016.
59. Fikkert, W.; van der Vet, P.; van der Veer, G.; Nijholt, A. Gestures for large display control. In Proceedings of the GW 2009: Gesture in Embodied Communication and Human-Computer Interaction, Bielefeld, Germany, 25–27 February 2009; pp. 245–256.
60. Valdes, C.; Eastman, D.; Grote, C.; Thatte, S.; Shaer, O.; Mazalek, A.; Ullmer, B.; Konkel, M.K. Exploring the design space of gestural interaction with active tokens through user-defined gestures. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, ON, Canada, 26 April–1 May 2014; ACM: New York, NY, USA, 2014; pp. 4107–4116.

61. Walter, R.; Bailly, G.; Valkanova, N.; Müller, J. Cuenesics: Using mid-air gestures to select items on interactive public displays. In Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services—MobileHCI '14, Toronto, ON, Canada, 23–26 September 2014; ACM Press: New York, NY, USA, 2014; pp. 299–308.
62. Roveló Ruiz, G.A.; Vanacken, D.; Luyten, K.; Abad, F.; Camahort, E. Multi-viewer gesture-based interaction for omni-directional video. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems—CHI '14*; ACM Press: Toronto, ON, Canada, 2014; pp. 4077–4086.
63. Walter, R.; Bailly, G.; Müller, J. StrikeAPose: Revealing mid-air gestures on public displays. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 27 April–2 May 2013; ACM: New York, NY, USA, 2013; pp. 841–850.
64. Ballagas, R.; Borchers, J.; Rohs, M.; Sheridan, J.G. The smart phone: A ubiquitous input device. *IEEE Pervasive Comput.* **2006**, *5*, 70–77. [[CrossRef](#)]
65. Boring, S.; Jurmu, M.; Butz, A. Scroll, tilt or move it: Using mobile phones to continuously control pointers on large public displays. In Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7, Melbourne, Australia, 23–27 November 2009; ACM: New York, NY, USA, 2009; pp. 161–168.
66. Lucero, A.; Holopainen, J.; Jokela, T. MobiComics: Collaborative use of mobile phones and large displays for public expression. In Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services, San Francisco, CA, USA, 21–24 September 2012; ACM: New York, NY, USA, 2012; pp. 383–392.
67. Liang, H.N.; Williams, C.; Semegen, M.; Stuerzlinger, W.; Irani, P. User-defined surface+motion gestures for 3D manipulation of objects at a distance through a mobile device. In Proceedings of the 10th Asia Pacific Conference on Computer Human Interaction—APCHI '12, Matsue-city, Shimane, Japan, 28–31 August 2012; ACM Press: Matsue-city, Shimane, Japan, 2012; pp. 299–308.
68. Medrano, S.N.; Pfeiffer, M.; Kray, C. Enabling Remote Deictic Communication with Mobile Devices: An Elicitation Study. In Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI '17, Vienna, Austria, 4–7 September 2017; ACM: New York, NY, USA, 2017; p. 19.
69. Bjork, S.; Holopainen, J. *Patterns in Game Design (Game Development Series)*; Charles River Media: Needham Heights, MA, USA, 2004.
70. Ren, G.; Li, C.; O'Neill, E.; Willis, P. 3d freehand gestural navigation for interactive public displays. *IEEE Comput. Graph. Appl.* **2013**, *33*, 47–55.
71. Du, G.; Degbelo, A.; Kray, C.; Painho, M. Gestural interaction with 3D objects shown on public displays: An elicitation study. *Int. Des. Archit. J. IxD A* **2018**, *38*, 184–202.
72. Hutchins, E.L.; Hollan, J.D.; Norman, D.A. Direct manipulation interfaces. *Hum. Interact.* **1985**, *1*, 311–338. [[CrossRef](#)]
73. McNeill, D. *Hand and Mind: What Gestures Reveal about Thought*; University of Chicago Press: Chicago, IL, USA, 1992.
74. Findlater, L.; Lee, B.; Wobbrock, J. Beyond QWERTY: Augmenting touch screen keyboards with multi-touch gestures for non-alphanumeric input. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, TX, USA, 5–10 May 2012; ACM: New York, NY, USA, 2012; pp. 2679–2682.
75. Vatavu, R.D.; Wobbrock, J.O. Formalizing agreement analysis for elicitation studies: New measures, significance test, and toolkit. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015; ACM: New York, NY, USA, 2015; pp. 1325–1334.
76. Wobbrock, J.O.; Aung, H.H.; Rothrock, B.; Myers, B.A. Maximizing the guessability of symbolic input. In Proceedings of the CHI'05 Extended Abstracts on Human Factors in Computing Systems, Portland, OR, USA, 2–7 April 2005; ACM: New York, NY, USA, 2005; pp. 1869–1872.
77. Du, G.; Lohoff, L.; Krukar, J.; Mukhametov, S. Comparing two methods to overcome interaction blindness on public displays. In Proceedings of the 5th ACM International Symposium on Pervasive Displays, Oulu, Finland, 20–26 June 2016; ACM: New York, NY, USA, 2016; pp. 243–244.

78. Sorce, S.; Malizia, A.; Gentile, V.; Gentile, A. Touchless gestural interfaces for networked public displays: Overcoming interaction blindness and performing evaluations in-the-wild. In Proceedings of the 2015 Adjunct ACM International Joint Conference on Pervasive and Ubiquitous Computing and 2015 ACM International Symposium on Wearable Computers, Osaka, Japan, 7–11 September 2015; ACM: New York, NY, USA, 2015; pp. 789–790.
79. Röcker, C.; Magerkurth, C.; Hinske, S. Information security at large public displays. In *Handbook of Research on Social and Organizational Liabilities in Information Security*; Gupta, M., Sharman, R., Eds.; IGI Global: New York, NY, USA, 2008; pp. 471–492.
80. Alt, F.; Schneegaß, S.; Schmidt, A.; Müller, J.; Memarovic, N. How to evaluate public displays. In Proceedings of the 2012 International Symposium on Pervasive Displays, Porto, Portugal, 4–5 June 2012; ACM: New York, NY, USA, 2012; p. 17.
81. Baldauf, M.; Lasinger, K.; Fröhlich, P. Private Public Screens: Detached Multi-user Interaction with Large Displays Through Mobile Augmented Reality. In Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia, MUM '12, Ulm, Germany, 4–6 December 2012; ACM: New York, NY, USA, 2012; p. 27.
82. Du, G.; Kray, C.; Degbelo, A. Interactive immersive public displays as facilitators for deeper participation in urban planning. *Int. J. Hum.-Comput. Interact.* **2019**. [[CrossRef](#)]
83. Wu, M.; Shen, C.; Ryall, K.; Forlines, C.; Balakrishnan, R. Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. In Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems, TableTop, Adelaide, Australia, 5–7 January 2006; IEEE: Piscataway, NJ, USA, 2006; pp. 183–190.
84. Barth, K.; Muller, W. Interacting with public displays for informal learning: Design issues and first experiences. In Proceedings of the 2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT), Timisoara, Romania, 3–7 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 92–94.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).