

Article

Designing Gestures for Data Exploration with Public Displays via Identification Studies

Adina Friedman *  and Francesco Cafaro * 

School of Informatics and Computing, Indiana University-Purdue University Indianapolis,
Indianapolis, IN 46202, USA

* Correspondence: adifried@iu.edu (A.F.); fcafar@iu.edu (F.C.)

Abstract: In-lab elicitation studies inform the design of gestures by having the participants suggest actions to activate the system functions. Conversely, crowd-sourced identification studies follow the opposite path, asking the users to associate the control actions with functions. Identification studies have been used to validate the gestures produced by elicitation studies, but not to design interactive systems. In this paper, we show that identification studies can be combined with in situ observations to design the gestures for data exploration with public displays. To illustrate this method, we developed two versions of a gesture-controlled system for data exploration with 368 users: one designed through an elicitation study, and one designed through in situ observations followed by an identification study. Our results show that the users discovered the majority of the gestures with similar accuracy across the two prototypes. Additionally, the in situ approach enabled the direct recruitment of target users, and the crowd-sourced approach typical of identification studies expedited the design process.

Keywords: elicitation studies; identification studies; embodied interaction; mid-air gestures; human-data interaction; data visualization



Citation: Friedman, A.; Cafaro, C. Designing Gestures for Data Exploration with Public Displays via Identification Studies. *Information* **2024**, *15*, 292. <https://doi.org/10.3390/info15060292>

Academic Editor: Danilo Avola

Received: 22 April 2024

Revised: 13 May 2024

Accepted: 16 May 2024

Published: 21 May 2024



Copyright: © 2024 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

End-user elicitation studies currently represent a state-of-the-art approach for designing interactive systems that do not have an established interaction vocabulary [1]. When conducting elicitation studies, the interaction designers prompt the participants with a system function and ask them what interaction command they would use to activate it. However, they face limitations: recruiting participants for in-lab studies may be a long, expensive, and cumbersome process [2–4].

Identification studies may mitigate some of these problems. Originally introduced by Ali et al. [4] to compare the gesture sets generated by multiple elicitation studies, identification studies reverse the traditional elicitation procedure: the users are prompted with an interaction command (in Ali et al.'s case study, a voice command [4]) and asked what system function it should activate. Interestingly, identification studies can be conducted through internet surveys, allowing the designers to reach a wider audience than what would be afforded in a traditional lab setting [4].

Both elicitation and identification studies are valuable methods for involving the users in the design process. They can be particularly beneficial in the context of embodied interaction [5–7], especially when designing gestures and body movements to control interactive data visualizations on public displays [8]. In fact, due to the limited availability of the established gestures and bodily movements [9] for data exploration, designers need to quickly create tailored gestures for every novel application.

In the interaction design literature, substantial attention has been directed to elicitation studies [10,11]. Conversely, the use of identification studies is newer. In this paper, we discuss how identification studies can be used in conjunction with in situ observations

to design embodied controls for data exploration with interactive public displays as an alternative to elicitation studies [12,13].

We structured our paper on the premise that “the best way to talk about methods is to show instances of the actual work” [14]. We conducted a study that developed two versions of a full-body gesture-controlled system for data exploration with 368 users: one designed through an elicitation study, and one designed through in situ observations followed by an identification study. We followed a three-stage approach to compare the effectiveness of the identification studies vs. elicitation studies (Figure 1).

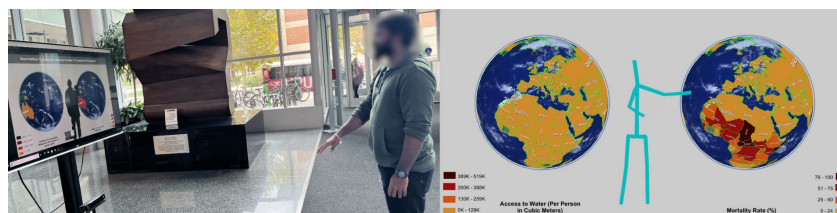


Figure 1. Left: One participant interacts with a public display in the campus center of a major urban university. Right: Detail of the data visualization on the display. It includes two data-sets shown side-by-side on a globe, and a representation of the user (in this case, as a stick figure).

In our study, we found that the participants were proficient in using both systems, successfully discovering the majority of the gestures required to activate the system functions. In summary, whereas the prior work used identification studies to evaluate the results of elicitation studies [4,15], we show that identification studies can be used as an alternative to elicitation studies to craft gesture-based interfaces, enabling the interaction designers to harness the advantages of in situ observations, as well as the cost-effectiveness and speed [16] offered by crowd-sourcing platforms.

2. Background: Physical Interaction with Data Visualizations

2.1. Embodied Interaction

Our work is on embodied interaction [5]. According to Dourish, we construct meaning through our embodied (i.e., physical, and situated in time and space) interaction with the world [5]. Hornecker (e.g., [6]) highlights the role of the user’s body and argues that “movement and perception are tightly coupled”. This perspective highlights the impact of physical movement on how users perceive, interact with, and make sense of the technology around them.

In this paper, we use the words “embodied interaction” to refer to interactive displays that are controlled by hand gestures and body movements (similar to the definitions used in [6,17]). In this context, embodied interaction represents a departure from traditional input devices like a keyboard and mouse: by utilizing gestures and body movements, the users can manipulate the interfaces and access information.

2.2. Human–Data Interaction (HDI)

Our work is on the design of gestures and body movements that people can use to interact with data visualizations on public displays. As such, it contributes to Human–Data Interaction (HDI) [8,18,19]. With HDI, we refer to a research stream investigating how embodied interaction [6] can facilitate people’s exploration of interactive data visualizations [18–20]. Rather than relying solely on traditional input devices such as a keyboard and mouse, HDI emphasizes the use of gestures and body movements as a means to manipulate, navigate, and make sense of complex data visualizations.

Because there is not an established interaction vocabulary for HDI [8], elicitation studies are currently the best approach to design the gestures and body movements in this context.

The reader should notice that, as highlighted by Victorelli et al. in their literature review of Human–Data Interaction [21], the term HDI has been used to refer to a broad range of research topics spanning from computer graphics to information science. Notably,

Mortier et al. [22] defined HDI as research on how people engage with “personal” data. Unlike Mortier, the work in this paper does not focus on personal data; rather, we adopted Elmqvist’s [18] and Cafaro’s [8,19] definition of HDI.

3. Related Work

3.1. End-User Elicitation Studies

End-user elicitation studies were introduced by Wobbrock et al. and allow designers to collect the user preferences for the symbolic input to control digital interfaces [1,10,11,23]. In these studies, the users are shown referents (the effect of their interaction) and asked to supply a symbol (the action that will lead to that effect) [23,24]. This allows the researchers to collect the symbols most preferred by their potential user base; the overarching idea is that such symbols are preferable to those created by HCI professionals alone [24]. However, the traditional in-lab elicitation studies face limitations related to their sample populations; these are often composed of potential users available on university campuses, which may not be representative of a wide range of users [4,15]. To overcome these limitations, Ali et al. [4] proposed conducting elicitation studies using a custom-made crowd-sourcing platform and validated this idea on a set of user-generated voice-based commands. Similarly, Gelicit presents a platform to conduct distributed elicitation studies over the internet by recording the gestures the users proposed for system functions and helping the researchers to analyze them [25], replicating the traditional elicitation study in a distributed form. These works greatly inspired the research that we present in this paper; the focus, however, was on multi-modal interaction including voice commands [4], while our work explores mid-air gestures and full-body movements to design the embodied interactions for data exploration.

3.2. End-User Identification Studies

Crowdlicit by Ali et al. [4] introduced a technique called end-user identification studies, in which the participants are asked to match the symbolic input to referents, reversing the process of an elicitation study, illustrated in Figure 2 [4,10]. Identification studies can be conducted online, which addresses some of the limitations of elicitation studies regarding the available subjects and resources needed [4,15]. They do come with their own limitations compared to elicitation studies, such as a lack of physical presence and performance of the gestures in front of the interface, which we address further in this paper. Previously, they were only used to evaluate the user-generated gesture sets created with elicitation studies [4,15]. In this paper, we build on this idea and show how identification studies can be combined with in situ observations and used as a design method for gesture-based, interactive data visualizations.

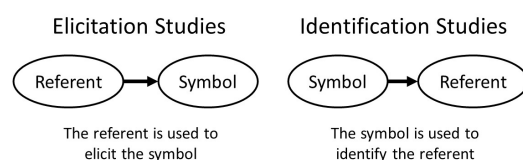


Figure 2. Illustration of the correspondences between elicitation and identification studies.

3.3. Mechanical Turk and Crowd-Sourcing

Although the Crowdlicit platform is no longer online, identification studies may be conducted through existing commercial platforms like Amazon’s Mechanical Turk (mTurk), which allows researchers (“requesters” in the mTurk system) to distribute surveys digitally through the platform and recruit participants (“workers”) [26,27]. The subjects recruited through mTurk can provide a sample population that represents the general population better than those recruited for traditional lab studies [26,28]. It should be acknowledged though that these platforms have issues, such as asymmetries between requesters and workers [27], and problems related to workers’ exploitation that need to be addressed or,

at the very least, mitigated [29]. However, the responses on Mechanical Turk have been found to be consistent over time and able to provide meaningful data for experiments that do not require specialized subject pools or social interaction [28].

3.4. Interactive Public Displays and User Representation

Interactive public displays are seeing increased use in public life in areas like museums [17,20,30], advertising [11,31], interactive art [32], or to deliver instruction content in locations such as airports or public forums [33]. The users can control the content using mid-air gestures, so they do not need to directly touch the screen and can interact from a distance that allows them to see the entire display [20,33,34].

People, however, may not notice the display, especially in highly stimulating public spaces, like city squares, museums, and community centers [35]. This phenomenon is known as display blindness [36]. Fortunately, integrating a representation of the user on the display can mitigate this design challenge. Tomitsch et al. [37] found that the users were sometimes more interested in playful interactions with their reflection on the screen, and system responsiveness to this sort of playful interaction could lead to greater user engagement. The previous research has explored which visual representations attract users to an interactive display [38], how they change their engagement with the display [33], and which representations help the users to identify themselves in a group [39]. There has also been work examining how to create gestures that are easier for a user to learn so they can more effectively interact with the system [40], including which gestures users perform based on the way they are represented while engaged in data exploration and browsing tasks [20,33].

Differently from Ali et al.'s [4] work, which considers a domestic scenario (people using their TV [24]), our work is about the design of public interactive displays. Thus, it must build upon the aforementioned literature: if we want passersby to notice and actually use the data visualization on the display, we need to integrate the user's representation on the screen.

4. General Methodology

Elicitation studies are frequently used to design gesture-based interfaces [4,10,20,40]. In these studies, users are provided a referent (typically something that the system can do) and asked to create or choose a symbol (an action/gesture/body movement to control that function). In contrast, identification studies provide the participant a symbol and ask them to choose which referent it should correspond to. In this experiment, the symbols are mid-air gestures/body movements conducted toward the display, and the referents are the system functions (e.g., zooming in).

4.1. Overview of the Stages

Our methodology uses two design stages, plus a third stage to evaluate the results obtained from the identification study against those from a traditional elicitation study (see Figure 3). During Stage 1, we collected gestures that users performed spontaneously in situ, in front of a public display that reflected their movements. We used a 65" screen with a globe-based data visualization (see header image). The in situ deployment allowed us to reach a large population of potential users faster than in a traditional lab study.

Based on the initial pool of gestures that we observed, in Stage 2 we conducted a crowd-sourced, survey-based identification study. The crowd-sourced approach allows designers to tap into a vast and diverse pool of participants and to quickly yield a large sample size, which is often challenging to achieve through traditional in-lab elicitation studies; recruiting using crowd-sourcing platforms is significantly faster and cheaper than gathering participants for lab studies [41].

Finally, Stage 3 compared the gestures and body movements crafted during Stages 1 and 2, vs. others designed using an in-lab elicitation study, to evaluate how effectively first-time users could "discover" them. We focused on discoverability because, in public

places, people cannot consult user manuals to learn which gestures and body movements they can use [17] and frequently leave thinking that the system is broken if the screen does not quickly respond to their attempts to interact [35].

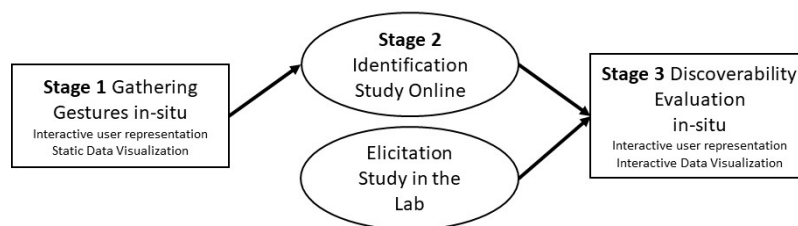


Figure 3. Illustration of the 3 stages of our study. Boxes are field deployments; circles are stages conducted in-lab or online.

4.2. Participants

Three-hundred-sixty-eight people participated in this study. Our activities were organized in three stages. Stage 1 consisted of 133 participants in the campus center. The survey in Stage 2 collected 104 online replies from participants in the US. The elicitation study collected input from 13 in-lab participants at the same urban university in the US. Stage 3 consisted of 118 participants recruited at the same campus center as Stage 1. This is shown in Table 1.

Institutional Review Board (IRB) permission was obtained to collect video recordings of in-person participants, and signs were placed by the display informing participants that they were being recorded. Because the observations in Stages 1 and 3 took place in situ and we wanted participants to be able to approach and leave the display at will, we did not stop them to conduct interviews or surveys. This prevented us from collecting additional demographic information about these participants.

Table 1. Participants in each stage.

Stage	Number of Participants
Stage 1 (in situ)	133
Stage 2 (online)	104
Stage 3 (elicitation study)	13
Stage 3 (in situ)	118

4.3. Implementation

4.3.1. Hardware Description

The system runs on an Intel® Core™ i7-4710HQ CPU @ 2.50 GHz 2501 MHz, 4 core(s), 8 logical processor(s) 16.0 GB RAM, and NVIDIA GeForce GTX 970 GPU. For the experiment that we describe in this paper, we used the Microsoft Kinect v2. The visualization was shown on a 65" TV screen.

4.3.2. Software Description

The main platform for the system and the on-screen visuals was the Unity3D 3D engine. We used the Kinect for Windows Software Development Kit (SDK) 2.0 for the Kinect camera for body tracking on the user representations, and developed the keyboard controls for the globes in C#.

4.3.3. User Representations

We implemented four different user representations or “mode types” [20], which mirrored the movements of users standing in front of the display. The four mode types were (1) stick figure—the user is shown as a stick figure that follows their movements,

(2) avatar—a jointed 3D figure that follows the user’s movements, (3) silhouette—the user is shown as a black silhouette, and (4) camera—the live video feed from the camera is shown with the background erased. The four mode types are shown in Figure 4.

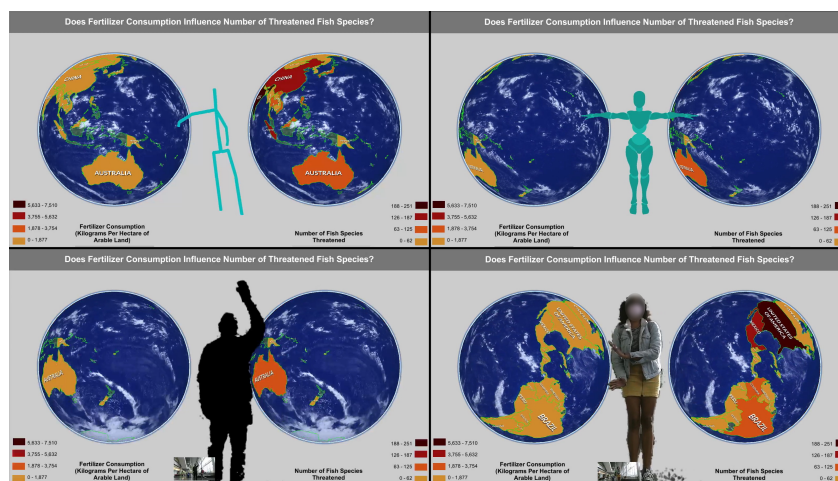


Figure 4. Screen captures of each user representation during Stages 1 and 3 (clockwise from top left: stick figure, avatar, full camera, and silhouette). During Stage 1, the globes in the background were not interactive.

4.3.4. Globe Visualizations

We created interactive globe visualizations by modifying a globe map package for Unity3D to provide realistic Earth and atmosphere settings that showed the boundaries of each country. The datasets used are automatically uploaded to the designated location on the map, and a gradient color is applied to each country reflecting the values loaded from the dataset files. The color values were normalized in the range [0–1] to create a consistent gradient scheme.

We chose to use this style of visualization as it has been successfully used before to facilitate casual data exploration in informal learning settings [20].

For Stage 3 of our experiment, the globes had simulated interactivity using a Wizard-of-Oz technique. A researcher was located near the screen with a computer where they could watch both the movements of the participants and the on-screen movements of the user representations. Although the “wizard” was within view of the participants, they did not actively draw attention to themselves, and moderators did not point them out.

5. Stage 1—Collecting Candidate Gestures and Body Movements In Situ, from Actual Users

The previous work on identification studies does not provide any guidance on how to generate the symbols (in our case, gestures and body movements) that the participants are exposed to. Ali et al.’s work [4] does not introduce identification studies as a self-standing design method; rather, they were conceived as a way to compare the symbols generated by different elicitation studies. Consequently, the symbols people were exposed to were those crafted during a prior elicitation study.

In order to enable interaction designers to use identification studies as a design method, we first needed to establish a procedure to craft an initial list of gestures and body movements. We opted for in situ observations. In Stage 2, the participants in the identification study are then shown symbols from this list and asked to match each symbol to a referent (i.e., a system function).

5.1. In Situ Observations: Procedure

Because our goal was to avoid the setting and population limitations of in-lab studies, we conducted this generative phase in situ. We set up an interactive display in the campus

center of an urban university in the US for two days to collect gestures in the intended context of use [11,20]. The campus center building is open to the public and is a social space, meaning that it lacks the formality of an in-lab study.

We introduced some interactivity elements to avoid the risk of display blindness [36]. An interactive representation of the user is ideal in this context because it has been found to attract users toward an interactive display [38]. In our study, passersby freely engaged with the display by standing in front of it and seeing their movements reflected on the screen by different visual representations of themselves, which a Microsoft Kinect enabled us to conduct in real time. These representations had a progressing degree of realism: a stick figure, 3D avatar, silhouette, and camera with background removed (see Figure 4). These user representations are known in the literature as “mode types” [20] and have been found to lure passersby toward interactive displays [20,38]. We used an exploratory approach and tested multiple “mode types” because the previous literature suggested that different ways of representing the users on the screen may lead to people performing different gestures [20]. Thus, we wanted to explore if some mode types were more informative than others for collecting gestures and body movements in situ. Using a quasi-experimental approach, these four user visualizations were displayed in a random sequence for 15 min each, with the users able to approach or leave the display at any time.

Importantly, during this phase, the data visualization (i.e., the two globes in Figure 4) was not interactive, and the participants were not informed of what future interactivity would be possible. We did not want to confuse the users by recognizing some gestures and not others; rather, we wanted to collect gestures and body movements that people spontaneously performed in front of the display and match them in Stage 2 with the functions of the data visualization.

We want to highlight that this approach enabled us to collect more data points over a shorter period of time than with an in-lab elicitation study: we observed 133 participants in two days (it took us two weeks to recruit 13 in-lab participants in Stage 2). Additionally, conducting Stage 1 in situ (rather than in the lab) mitigated some of the problems of using surrogate users [3] who may not be representative of the target population [4].

5.2. In Situ Observations: Analysis

We recorded videos of the participants interacting with the display, then adopted an approach based on Interaction Analysis [14]. Using the VGG Image Annotator [42], four researchers collaboratively coded the videos working in pairs with an initial set of gestures created through all four researchers reviewing part of the footage as a group. All the gestures from this initial session and subsequent ones were added to a coding dictionary that was referenced and expanded as the research progressed, with some earlier videos being re-coded to include newer gestures. During this process, the researchers grouped and reviewed instances of gestures and body movements that they deemed as “substantially similar” [43] (i.e., slight variations of the same gestures). The full team of researchers reviewed and discussed instances of disagreement during two 2-hour meetings using the shared dictionary to resolve any disagreements between the researchers. The resulting data included the start time and identifying name for each gesture, as well as the mode type [20] in which those gestures appeared.

This provided us a pool of gestures that passersby spontaneously performed in front of our display in the intended context of use. The full dictionary of gestures that we identified is included in Appendix A.

In contrast to previous elicitation studies such as those conducted by Morris [24], our gestures were collected in situ and without showing users a direct functional correspondence to their gestures. Thus, it does not require a functional prototype (our data visualization was intentionally non-interactive in Stage 1), nor to actively recruit participants (we simply observed passersby).

5.3. Results: Common Gestures Performed toward the Display

We collected a total 667 data points in Stage 1 (because the study was conducted in situ, with no direct supervision, passersby performed as many gestures as they liked). Overall, the 133 people in Stage 1 performed a total of 47 distinct gestures and body movements (we grouped the individual data points into 47 gestures using the coding approach described above).

In line with the findings in [20], there were differences in the gestures performed for each type of user representation (“mode types” [20]). We observed, however, that there were gestures that appeared across all the types of user representation. Interestingly, we noticed that some gestures appeared in the top five most-performed gestures across mode types (eliminating idle arm movements): hand wave one hand, arm wave one arm, dancing, kicking, and arm wave two arms.

This indicates that the way users are represented on the screen does not always impact the gestures they use to interact with a display: there are some gestures that are common across all mode types, i.e., gestures that people spontaneously perform when they are in front of any of the mode types listed in [20].

Additionally, we coded the gestures by their locus—which we defined as which part of the body was moved while performing the gesture. Over 50% of all the gestures performed had a locus in the hands or arms (34.04% in the arms and 23.98% in the hands). This indicates that our users prefer moving their upper bodies to interact with a display, which is in line with the findings from Narvaes et al. that many elicitation studies resulted in gestures performed with the hands or arms [10].

6. Stage 2—Matching Gestures/Body Movements to Functions of the Data Visualization Using Identification Studies

6.1. Identification Study: Procedure

In Stage 2, we conducted an identification study using an approach similar to the one introduced in [4]. We created a survey (Figure 5) using Qualtrics, in which we presented participants with six functions of the data visualization: rotate up, rotate down, rotate clockwise, rotate counter-clockwise, zoom in and out, and switch dataset. These functions are based on the existing literature regarding how to support data exploration using interactive data visualization in public spaces (e.g., see [17]).

To narrow down the list of gestures from those listed in Appendix A, we took the most-performed gestures that only required one person to perform and were consistently repeatable. This eliminated gestures such as high-fiving, which requires two people, and gestures such as those coded as ‘exploratory hand movements’ and ‘idle arm movements’, which could not be repeated just from the coding description. Gestures that could be performed with a single limb were divided into left and right sides of the body.

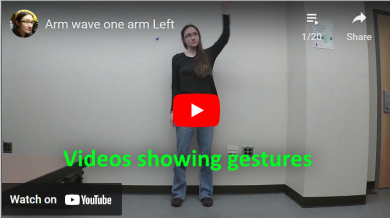
On the right side of the survey screen, the participants were provided a list of gesture options and asked to match a single gesture to each of the system functions. To avoid ordering effects, the order of the gesture options was counterbalanced. The participants were limited to those in the US, and other demographic data were not collected. We did not ask the participants to explain the reasons behind their choices.

The survey was distributed through Amazon’s Mechanical Turk service, with 120 surveys distributed in two 60-person batches. The researchers checked back after 24 h for completion of the surveys, at which point they were closed. In all the cases, we were able to obtain 60 responses within the 24 h period.

We relied on the Amazon’s Mechanical Turk crowd-sourcing platform, which allowed us to collect survey responses quickly. To help improve workers’ trust in the research team (a known challenge with mTurk is the issue of trust between the requester and the workers [27]), we included our contact details as the ones conducting the study, as well as providing the means to contact our institution’s IRB. For the purpose of the identification study, we were not specifically concerned with the validity of the data as previous research has shown that data gathered from participants on the internet are not poorer-quality

than those collected from subjects by other means [26]. Additionally, due to mTurk's requirement that each individual's user ID be connected to a single bank account, the chances of repetition are decreased.

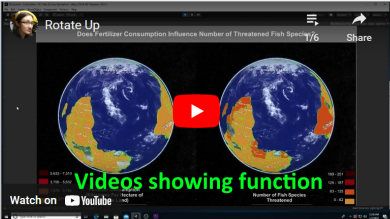
See this playlist for demonstrations of each gesture if the names are unclear.



Watch on YouTube

Please select a unique gesture for each function, videos demonstrating each gesture are available at the top of the page

This playlist contains the prior demonstrations of System Functions



Watch on YouTube

Select a different Gesture for each System Function
note: Left and Right versions are different gestures

Items	
Arm wave Left arm	Rotate Globe Up
Arm wave Right arm	
Arm wave two arms	Rotate Globe Down
Background push	
Background try to click	
Doing the wave	
Grab and move	
Hand circles	Rotate Globe Clockwise
Hand wave Left hand	
Hand wave Right hand	
Hands in and out	Rotate Globe Counterclockwise
Kicking	
Leaning Left	
Leaning Right	
Raising both arms	
Reach Left	
Reach Right	
Swipe Left to Right	
Swipe Right to Left	
Walking side to side	
	Zoom in and out
	Switch Dataset

Drag and drop gestures to match functions

Figure 5. The survey used to conduct the identification study, with the video playlists on the left and the drag and drop matching on the right.

6.2. Identification Study: Analysis

We evaluated 104 responses to the survey out of 120, with 16 replies eliminated due to incomplete or unusable responses. We took the most commonly selected gesture or gestures for each function and assigned them to control that function.

For gestures where the specific side of the body used did not directly relate to the on-screen visual (i.e., moving elements on the right or left side of the screen with the right or left hand), the gesture was generalized. This was the case in rotating the globe up for the identification study, in which 'Arm wave right arm' was the most selected, with 'Arm wave left arm' also receiving a substantial number of votes, so both sides were used to activate the rotate up function—an approach that can aid left-handed users [44].

6.3. Results: Control Patterns and Trends in Replies

Table 2 reports the set of gestures that were crafted at the end of the identification study procedure.

The results of the identification study assigned four out of the six functions to upper-body movements. Two exceptions were rotate down being assigned to kick and change dataset being assigned to walking side to side. This is in line with our observation from Stage 1 that people (at least our user population) tend to use their arms and hands to interact with a display more than they use their lower bodies or whole bodies.

When two gestures of body movements were recommended by the same number of people, we included both of them in the user-generated set of gestures rather than selecting a single one. For instance, left hand wave and swipe right to left were both included to

control rotate globe clockwise in the identification study. Notably, this only happened in the identification study as there was always a clear consensus in the elicitation study. Clockwise rotation ultimately had the most votes for waving the corresponding hand, while counter-clockwise rotation had the most votes for swiping in the corresponding direction. To preserve internal consistency [45] when rotating the globe either clockwise or counter-clockwise, we combined these results and allowed either hand waving or swiping to be used for the rotational function. Both of these gestures have the final location of the hand placed on the side of the body corresponding to the way they want the globe to rotate. Zooming in and out had an equal number of votes for ‘spreading hands, bringing hands together’ and for ‘swipe left to right’. However, since the swiping gestures had more votes assigning them to the rotational functions, and one gesture could not control multiple functions, ‘spreading hands, bringing hands together’ was selected for the zoom functions.

Table 2. Control patterns that were crafted from the identification study and from the elicitation study

System Function	Identification Study	Elicitation Study
Rotate the Globes Up	Arm wave (above waist)/Raising both arms	Swipe Up
Rotate the Globes Down	Kick	Swipe Down
Move Globes Clockwise	Left hand wave/Swipe right to left	Swipe right to left
Move Globes Counter-Clockwise	Right hand wave/Swipe left to right	Swipe left to right
Zoom in and out	Spreading hands, bringing hands together	Spreading hands, bringing hands together
Switch Dataset	Walking side to side	Clicking/Pointing/ Pushing

7. Stage 3—Evaluation

To evaluate the two-stage method that we introduce in this paper (in situ observations followed by identification study), we conducted a separate elicitation study to design the gestures for our data visualization [1]. Elicitation studies are a well-defined and established method for creating control patterns for gesture-controlled systems, thus providing us a reliable benchmark against which to test our method [10].

Importantly, when using the in situ observations plus identification study approach that we introduce in this paper, researchers and practitioners do not need Stage 3; it is not part of the procedure that we outline for identification studies. Its only purpose is to evaluate our approach. Previously, identification studies have been used to verify the results of elicitation studies [4]; in Stage 3, we reversed this process, using the elicitation study to test the results of the identification study.

7.1. Step 1—Elicitation Study: In-Lab Procedure and Resulting Gesture Set

During the in-lab elicitation study, the participants were shown pre-recorded animations of each of the system functions and asked by a moderator what gesture or body movement they would perform to activate it. To ensure that the gestures were recorded correctly, the moderator would ask clarifying questions after some movements, such as “To be sure, swiping right to left would rotate the globes clockwise and swiping left to right would rotate the globes counter clockwise?” and check for the participant’s agreement. We collected video recordings from 13 participants that we recruited at an urban US university campus over the course of two weeks. These videos were analyzed by two researchers, cross-checking to make sure they agreed on the gesture descriptions.

Gesture Set from the In-Lab Elicitation Study

Table 2 reports the set of gestures crafted with the in-lab elicitation study. Interestingly, some of these gestures were different than those generated with the identification study (e.g., rotate up), while others were identical (e.g., zooming in and out).

7.2. Step 2—Discoverability Evaluation (In Situ)

To assess how the gestures designed with in situ observations plus identification studies compare with those crafted with a traditional elicitation approach, we focused on the discoverability [30,46] of the resulting gesture set. Specifically, we wanted to see if the users were able to guess the gestures that our prototype was able to recognize. To preserve the ecological validity [47] of this evaluation, we did not include any scaffolding on the screen or instructions from the moderator (the gestures and body movements were designed for interactive visualizations on public displays).

7.2.1. Procedure

We once again set up an interactive display at the same campus center of an urban university in the US, showing the same four modes of representing the user that we used in Stage 1: a stick figure, 3D avatar, silhouette, and camera (Figure 4). These representations were shown in front of an interactive data visualization showing two globes that moved together in response to the user's movements (see Figure 6). Unlike in Stage 1, which had an interactive user representation and static globes, this display had an interactive user representation and interactive globes. While the figures used a Microsoft Kinect camera to track and reflect the user movements, interaction with the globes was accomplished using a Wizard-of-Oz technique [48]: we wanted to accurately assess if the users were able to guess our gestures/body movements, not the accuracy of the gesture recognition system.

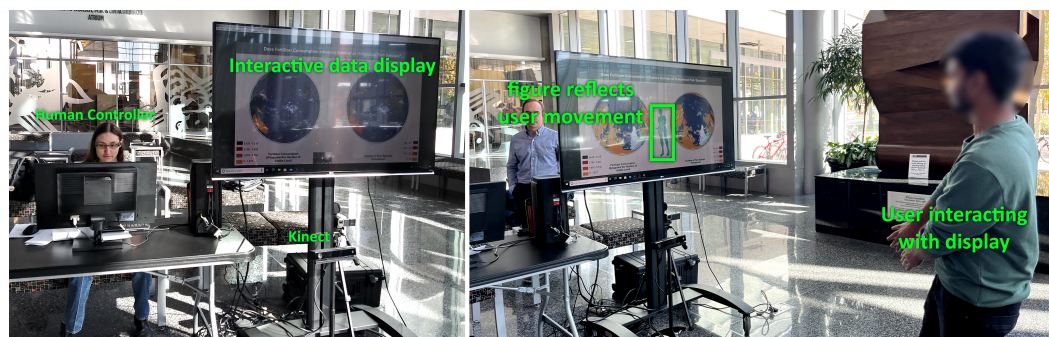


Figure 6. Left: The setup used to test the different control patterns. Right: Example of participant interaction with the display.

The field test was conducted over two days. As is common practice for in situ research in public spaces [30], we used a quasi-experimental design: the participants were not randomly assigned to an experimental condition but interacted with the version of the system that was active at the time of their visit. Each condition (gestures from identification/elicitation study + user representation) was shown for 15 min before moving to the next. The order of conditions was randomized and different on both days.

Differently from Stage 1, the participants in Stage 3 were actively recruited by a moderator among the passersby in the campus center. Additionally, while the system we used for Stage 1 supported multiple users, representing each one with an individual avatar or stick figure, the system for Stage 3 allowed only one user to interact with and control the system at a time. This decision was made to narrow down the range of gestures so single participants taking a survey could more easily envision themselves using the gestures (this better aligned the identification study in Stage 2 with a typical elicitation study, in which users are interviewed one at a time).

7.2.2. Analysis

Screen recordings of participants' interactions were analyzed by six researchers working in pairs to find how successfully the users interacted with the visualization in each control pattern. These videos were evaluated in terms of hit or miss, hit meaning that a user activated a function by performing the correct gesture (i.e., spreading their hands apart and seeing the visualization zoom in) and miss meaning that they either never discovered a function and/or never performed the correct gesture to activate it.

7.3. Results: Comparing Timing and Number of Users

Overall, the elicitation study took place over the course of 2 weeks, with the researchers actively trying to recruit participants. In contrast, the identification study was conducted within 24 h over the internet with a higher response rate (104 valid responses). Additionally, although it only took place over 3 days, the in situ observations in Stage 1 gathered more gestures from a greater number of participants (13 in the elicitation study vs. 133 during the in situ observations).

7.4. Results: Discoverability of Gestures and Body Movements

The passersby were free to begin interacting with the display and leave at any time, and spent on average 1 min and 22 s interacting with the display, regardless of the control pattern or mode type. This aligns with the previous findings showing that people spend on average around 1 to 2 min with an interactive installation in public spaces [49].

7.4.1. Effect of the Design Method (Identification vs. Elicitation)

The control patterns (see Table 2) were evaluated in terms of hit or miss. A system function was hit if a participant managed to activate it (i.e., if the participant was able to discover the gesture or body movement to activate that function), while a function was missed if the participant never activated it. Table 3 shows the percent of participants that were able to activate each system function in the two experimental conditions.

Table 3. The percentage of participants who hit each of the functions by performing the gesture needed to activate it. * beside the system function indicates a statistically significant difference.

System Function	Participants Who Discovered Gestures from Identification Study	Participants Who Discovered Gestures from Elicitation Study
Rotate Globes Up	91.67%	93.18%
Rotate Globes Down *	25%	81.82%
Move Globes Clockwise *	94.44%	75%
Move Globes Counter-Clockwise	86.11%	70.4%
Zoom In and Out	50%	54.55%
Switch dataset	43.67%	25%

To assess if there was a statistically significant relationship between the design method (elicitation vs. identification study) and the movements that the participants were able to discover, we created a table in which we listed, for each participant, the design method, and whether they were able to guess the gesture for each of the six system functions that we listed in Table 2. For example, we coded P1 as 1 for “rotate up” and 0 for “switch dataset” because P1 was able to discover the proper gesture to control rotate up but not the one to control switching the dataset during the Wizard-of-Oz evaluation study. Because we wanted to assess if there was a relationship between a nominal variable (i.e., the design method) and six dichotomous variables (whether a participant discovered the gesture for rotate up or not, for rotate down or not, etc.), we used chi-squared tests of independence. Specifically, we performed chi-squared tests of independence to examine

the relation between the design method (elicitation vs. identification study) and whether the participants were able to discover the movement to control each function.

The relationship between the design method and whether the participants were able to guess the gesture for “rotate globes down” was significant, $\chi^2(1, N = 80) = 25.972$, $p < 0.001$. Eighty-one percent of the participants in the elicitation condition were able to guess the control gesture for “rotate down”, compared to 25% of the participants in the identification condition.

Additionally, the relationship between the design method and whether the participants were able to guess the gesture for “rotate globes clockwise” was significant, $\chi^2(1, N = 80) = 5.501$, $p < 0.019$. Seventy-five percent of the participants in the elicitation condition were able to guess the control gesture for “clockwise”, compared to 94% of the participants in the identification condition.

There was no statistically significant difference for any other system function.

In other words, the two design methods (elicitation and identification studies) yielded comparable system results when considering the discoverability of the user-defined gesture set: one performed better for rotate down, the other for clockwise, and they had comparable results for all the other functions.

7.4.2. Effect of the User Representation (Mode Type)

We conducted a two-way ANOVA to assess if there was an effect of the design method (elicitation vs. identification) and mode type (whether the user was represented as an avatar, stick figure, silhouette, or full camera) on the total number of gestures that each participant was able to guess during the Wizard-of-Oz evaluation study. There was homogeneity of variance, as assessed by Levene’s test for equality of variances ($p = 0.328$). There was no statistically significant interaction between the effects of the design method and mode type on the average number of discovered gestures, $F(3, 72) = 0.555$, $p = 0.935$. In other words, the alternative ways of representing the users (mode types) that we considered in this study did not significantly alter the effect of the design method on the average number of gestures that the participants were able to discover.

8. Discussion

Identification studies have the advantage of being able to be performed remotely, allowing them to reach a wider audience than elicitation studies. Additionally, as we discuss in this section, using in situ observations plus an identification study provided insights that may mitigate the interaction and affordance blindness [50,51] when designing gestures and body movements to control the interactive data visualizations on public displays.

8.1. Interaction Blindness

People walking past a public display may not notice that the system is interactive: this is a problem known as “Interaction blindness” [50].

8.1.1. Wave Gestures and Mirrored Posed are Entry Points to the Interaction

Including discoverable entry points to the interaction (like wave gestures) is extremely important to engage people with large displays in public spaces [35], where the users cannot consult user manuals to learn how to interact with a display [17] and frequently leave thinking that the system is not interactive if it does not quickly respond to their actions.

During the in situ observation study, we noticed two usage patterns that can help to mitigate this problem.

First, the gesture that we observed the most across all the mode types was hand wave one hand. The participants would often approach the display and wave and then see this movement reflected by the figure on the screen. We believe this may have a social connotation: this could be viewed as the participants greeting the system, then interpreting the character on the screen “waving back” as the system returning the greeting. Starting

with a waving gesture may also be a result of legacy bias [24] as this sort of gesture is used to activate the Microsoft Kinect on older Xbox gaming systems. The prevalence of waving as a gesture to interact suggests that it may be a good entry point into interacting with the display because the participants see an immediate response to their actions.

Second, many participants approached the display and adopted a T-pose with their arms held at shoulder height and their feet together. This happened primarily with the stick figure and 3D avatar; in both conditions, the on-screen figures defaulted to a T-pose when they were not tracking a participant. As a result, the participants may have believed that they needed to imitate the on-screen pose to begin their interaction. This may also have come from unconscious mimicry and the human instinct to mirror someone with whom they are interacting [52]. Thus, carefully designing the starting pose of the user representation may also provide an entry point to the interaction by implicitly communicating to the user how to start operating the data visualization.

8.1.2. Multiple Gestures for the Same System Function

In the case of rotations, the identification study provided two gestures for them (hand waves and swiping), while the elicitation study provided one (swiping). As we mentioned in the Results section, this happened because the two gestures were recommended by exactly the same number of participants in the identification study. Similarly, the identification study also provided us with two different gestures to rotate the globe upward (arm wave (above waist) and raising both arms), while the elicitation study generated one (swiping upward).

The presence of multiple options in the identification study is likely attributed to its ability to engage a larger and more diverse range of participants, offering an increased number of perspectives and preferences. This finding warrants further investigation as it suggests that a broader participant base may contribute to the generation of multiple gestures for a single function. Ultimately, it could improve the discoverability of gestures: as observed in [35], using multiple gestures to control the same function provides additional entry points to the interaction.

8.2. *Affordance Blindness*

Even after they start interacting with public displays, the users may not be able to discover the gestures and body movements to activate all the system functions; this is a problem known as “affordance blindness” [51].

8.2.1. In Situ Approach and Legacy Biases

The participants in both design methods (identification and elicitation) chose spreading hands and bringing them together again for zoom in and zoom out, and swiping across the body to rotate the globes (with the identification study adding waving a hand in the direction the globe should turn). This consistent choice of gestures across both design methods implies a potential consensus on which gestures should trigger particular changes or actions on a visualization, particularly when the participants share a common geographical background, as was the case in this study, which focused on the US. Thus, including these recurrent gestures as a means to control interactive data visualizations can help to mitigate the affordance blindness [51] problem in public displays because they are consistent across the target user population.

These results may also be influenced by individuals’ prior experience with touchscreens, where gestures like swiping and pinching are commonly utilized. Legacy bias often permeates gesture elicitation studies as the users tend to suggest gestures they are already acquainted with [53,54]. In some application scenarios (especially when the participants are exposed to novel interactive systems they are not familiar with), legacy biases can derail the elicitation study. For example, the work in [30] describes the case of a pinch and zoom gesture (common regarding touchscreens) that was selected after an elicitation study to control the zoom function of a map-based data visualization. However, during

the real-world testing at a science museum, no one attempted to use this gesture. In other words, a gesture performed frequently during the in-lab elicitation study was never discovered during the in situ deployment of the system [30]. The gestures and body movements devised using the method outlined in this paper may be less prone to this issue since we observed user actions in real-world settings during Stage 1. However, it is essential to acknowledge that no research method is entirely impervious to potential biases, so this should be further investigated in future studies.

8.2.2. Upper- vs. Lower-Body Movements and the Role of Interaction Designers

The identification study condition performed significantly worse than the elicitation study condition in rotating the globe down. In this case, the identification study used kicking the feet, while the elicitation study used swiping down. This seems to indicate that, once the users initially see success interacting with the display using their upper body, they may not think to try using their legs to interact. This is consistent with the fact that upper-body movements, including the arms and hands, are often more prominently used for tasks that require fine motor skills, the manipulation of objects, and interactions with the environment [55]. For example, typing on a keyboard, writing with a pen, cooking, using a computer mouse, and gesturing while speaking [56] predominantly involve upper-body movements. In the identification study, the users were presented with the kick as an option, while they were not prompted to consider their lower bodies in the elicitation study. Thus, interaction designers need to carefully consider the gesture set generated with the in situ observations in Stage 1 before moving to Stage 2; when this set includes lower-body movements, it can be problematic in terms of discoverability.

Notably, the participants had fewer hits with the motions that required them to move their lower bodies, such as kicking to rotate the rotate globes down, or walking side to side to switch datasets. This seems to indicate that the users are accustomed to using their upper bodies to interact with the display and do not think of using their lower bodies. Interestingly, the walking side to side gesture could be discovered by accident as the participants approached or left the display, which may mean it is well-suited as an entry point to the interaction; alternatively, these sorts of accidental interactions could be used to better introduce functions that participants may not think to look for (we also highlight an issue with the participants not knowing to look for the switch dataset function, which may have impeded more intentional gestures like ‘clicking’). It may also indicate that our user population preferred to use gestures that only require them to move one or two limbs while staying in one spot rather than whole-body movements that would shift their position in front of the screen and could obstruct their view.

In our experiment, we focused on the user representation as an entry point to the interaction in Stage 1 rather than an interactive data visualization. Basically, we wanted to identify gestures and body movements that users can easily discover without any instructions or scaffolding. This may have resulted in gestures that may be more suited to interacting with the user representations rather than data exploration. In this vein, the addition of interactive data visualizations may have changed users’ attitudes toward the display, making them more focused on the data exploration and less playful. As a result, movements like dancing and kicking may not have been used as much when we switched the focus to the interactive data visualization. Interestingly, the social space in which we conducted Stages 1 and 3 was the same (the campus center), so this change in users’ preference cannot be interpreted as gestures that people feel shy to perform in a public space. Thus, interaction designers may once again need to select gestures that would be both good entry points and appropriate for data exploration.

The preference for upper-body movements that we observed, however, may also depend on the specific user population: the distribution of upper- and lower-body usage can be influenced by individual preferences and physical conditions. Intuitively, the in situ observations (Stage 1) may facilitate the selection task for interaction designers

because they provide direct insights on the user populations. Future work should validate this hypothesis.

8.3. Common Gestures Are Used Regardless of the User's Representation (Mode Type)

While the previous research has delved into the exploration of various user representations (mode types) in the contexts of interactive data visualizations [20] and public displays [33,37–39], our study yielded an unexpected finding: the type of user representation did not have a significant influence on the most frequently observed gestures. This mitigates concerns that manipulating the user representation might inadvertently impact the gestures used as entry points that facilitate the discoverability of system functions. It also indicates that different user representations can be used for other functions, such as driving engagement [20], without the concern that they will influence the controls needed to operate the interactive display.

8.4. When to Use Identification Studies

In this paper, we have presented an approach that combines in situ observations with identification studies as an alternative design method whose results are comparable to the traditional elicitation studies. Thus, elicitation and identification studies can be viewed as different approaches to achieving the same goal, depending on the resources available to designers (such as their access to potential users, spaces, and prior work in their application domain). Identification studies, when informed by in situ observations, require access to the physical space in which the system is intended to be deployed, and to a population of users. On the other hand, they do not require a lab or experimenters to administer the study.

Compared to elicitation studies, identification studies may require more work to be completed upfront: the collection of gestures through in situ observations, the creation of gesture pools, and the distribution of surveys. Thus, they are ideal when there is a need to collect data from many users in a short amount of time because the rapidity to collect data through in situ observations can quickly offset this trade-off. Furthermore, it is also easier to recruit participants online than in person. Depending on how they are distributed, identification studies can return more results, and faster than traditional elicitation studies (this was the case with both Crowdlcit [4] and our study). Our elicitation study was conducted over the course of 2 weeks with 13 participants, while our identification survey in Stage 2 took less than 48 h cumulatively to collect the data regarding over 100 participants.

While elicitation studies require the participant responses to be coded after the in-lab study is completed, with an identification study, the gestures have already been coded before the identification portion (in our paper, Stage 2) can be conducted. By pre-coding the gestures, identification studies enable a more streamlined data collection process, which can be helpful for applications when designers need to constrain the possible gestures used to activate the system functions. In our case, for example, we did not want the individuals to directly touch the screen, so it was not presented as an option in the identification study. Such limitations on gesture options can be more transparently communicated through an identification study, whereas an elicitation study might necessitate explicit statements regarding disallowed gestures.

We believe that the gesture pools identified during in situ observations (Stage 1) hold the potential for reuse across various displays within the same setting or location. For example, if we had to design multiple interactive data visualizations for the student center where we conducted our in situ observations, we could reuse the catalog of gestures in Appendix A. By leveraging previously identified gesture pools, the designers can bypass the initial stage of collecting and cataloging gestures, thereby further expediting the design process. The future work should investigate how and when it is acceptable to skip Stage 1 by using pre-collected catalogs of gestures.

Ultimately, the decision to use identification or elicitation studies may depend on the resources available to the designers.

9. Limitations

A limitation of identification studies is that they require an existing pool of gestures to act as options for the identification study. In our experiment, we used an initial in situ deployment to collect the gestures in context (Stage 1). However, we want to acknowledge that these gestures could be gathered or created in different ways. Researchers could use the gestures created in systems similar to the ones they are designing (for our case, these would be the ones found in [4,8,20,30,33]), or by consulting interaction designers to create a pool of gestures. In this vein, we have included our dictionary of gestures collected in Stage 1 in Appendix A and invite future researchers to make use of it if they see fit. These gesture pools may be limited by their intended context of use though (the gestures designed for casual data exploration may not be suited to in-depth data analysis), and future research could examine this assumption.

The lack of limitations regarding which gestures the users could perform in front of the display in Stage 1 may have resulted in some impractical gestures and sets of gestures that had no consistency or coherence. Additionally, the lack of animation in the globes may have influenced how the participants tried to interact with them before changing their focus to their on-screen representations. An interaction designer could help to filter these outliers from the pool of results. We were more focused on the entry points to the interaction with non-interactive globes.

Elicitation studies have the benefit of being able to ask the participants to physically perform the gestures they are suggesting, which may influence their choices. For example, in the identification study, the users chose kicking to activate the rotate down function; in the elicitation study, the users chose swiping downward, which was in line with their use of upper-body movements for all the other functions. The participants in an identification study can be asked to perform the gestures physically, but, without any sort of video verification, it is difficult to know if the participants actually do so. Future work should investigate how to address or mitigate this issue.

Not all system functions are the same. Further analysis of the system indicated that we may have an affordance blindness [51] issue with the option to switch the dataset on display. This function is not immediately obvious to users: while they can more easily expect to zoom in or out, or to rotate the globe, they may not know that they can explore different data. Thus, they may not think to look for this function when operating the system, and this may have contributed to the relatively low discoverability of that function (see Table 3). A future version of the system may have a scaffolding or signifiers [57] for this function, such as tiles showing which dataset the user has selected and text prompting them to explore the next set of data.

Our in situ studies were conducted on an urban university campus in a building open to the public because our final prototype was meant to be deployed in that space. This may have made our sample more representative of our target population. However, it may not accurately reflect the true distribution of the US population; that was not our goal (we wanted to design a system for that specific pool of users), and it would require a different method to collect the gestures and body movements in Stage 1. Additionally, although there were more logistical challenges to the in situ deployment (booking the location and transporting equipment and personnel), the amount of data that we were able to collect in a short period of time made these additional steps worthwhile.

10. Conclusions

In this paper, we describe a method to use identification studies combined with in situ observations to design gestures and body movements to control the interactive data visualizations on public displays. We compared how a prototype of an interactive data visualization performed with respect to a version of the system designed through a traditional in-lab elicitation study method. Overall, compared to the control gestures provided by the elicitation study, the identification study provided control gestures with a similar degree of discoverability for five out of the six display functions. This means that

interaction designers can create novel interfaces by leveraging the speed and population advantages of in situ observations and crowd-sourced identification studies. Thus, the identification study provided results comparable to the state-of-the-art elicitation studies.

Future work should explore the use of identification studies for designing collaborative gestures and interfaces, and further investigate how designers can actively refine the user-generated gesture sets to maximize the usability and discoverability of embodied interactions. The future work may also explore the different domains of interaction beyond interactive data visualizations.

Author Contributions: Conceptualization, A.F. and F.C.; methodology, A.F. and F.C.; software, A.F.; validation, A.F.; formal analysis, A.F. and F.C.; investigation, A.F. and F.C.; resources, F.C.; data curation, A.F.; writing—original draft preparation, A.F.; writing—review and editing, A.F. and F.C.; visualization, A.F.; supervision, F.C.; project administration, A.F. and F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of Indiana University (protocol code 12803, date of approval 22 December 2021; and protocol code 15461, date of approval 7 June 2023).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Dataset available on request from the authors. The movement dictionary is included in Appendix A.

Acknowledgments: We want to thank Aravindh Nagarajan for his contribution to the system implementation and Shweta Singh, Koushik Tripurari, Sachin Kawa, Hinal Jagdish Kiri, and Mary Kate Musholt for their help with the in situ study.

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

Abbreviations

The following abbreviations are used in this manuscript:

ANOVA	Analysis of Variance
HCI	Human-Computer Interaction
HDI	Human-Data Interaction
IRB	Institutional Review Board
MDPI	Multidisciplinary Digital Publishing Institute
SDK	Software Development Kit
U.S	United States

Appendix A. Movement Dictionary

We have compiled the gestures gathered in Stage 1 into a movement dictionary. This dictionary contains the names and descriptions of the gestures collected. We have included them because other researchers designing similar public displays may use some of these gestures that the users spontaneously performed in Stage 1 in the design of their systems. Additionally, it can serve as a starting dictionary for other designers and researchers who may want to use the in situ approach that we detailed in Stage 1.

Format of gesture definitions:

Label: description of the gesture, BODY PART

Arm circles: moving the arms/lower arms in circular movements, away or close to the body, ARM

Arm swing: small arm movements made close to the body—usually not showing clear intention or distinct beginning and end points, ARM

Arm wave one arm: waving using the entire arm (moving from the shoulder joint), more than 45 degrees, ARM

Arm wave two arms: waving both arms using the entire arm, moving from the shoulder joint (both at the same time or closely together), more than 45 degrees, ARM

ATC (air traffic control): gestures similar to ground crew gestures, usually with upper arms held at shoulder length while moving the forearms, ARM

Background flip: trying to grab content on the screen/background object and reversing hand positions/moving hands across each other

Background grab and move: tries to close their hand to “grab” content on the screen/background and move it

Background push: person is trying to push content on the screen/an object in the background, usually conducted by aligning a part of their body with the edge of an object and making a pushing motion

Background try to click: trying to interact with content on the screen/background to click or select an element, usually through a tapping motion performed in mid-air

Background Undefined: person is trying to interact with content on the screen background but their intention is not clear to observers

Clapping: strike the palms of the hands together repeatedly, typically in order to applaud someone or something, HAND

Dancing: Full-body coordinated movement, conducted purposefully in a rhythmic way, and can be recognized as a dance using common sense or previous designation, WHOLE BODY

Doing the wave: arm movements performed with the arms perpendicular to the body, showing a ‘wave’ by moving parts of the arms up and down sequentially from left to right or right to left, ARM

Error: person encounters a glitch and may need to walk in and out of detection to fix it; this should not count as deliberate behavior toward the display

Exploratory finger movements: moving the hands and fingers, mostly as an experimental gesture to see how the system would respond, moving fingers into different positions relative to the palm, HAND

Finger puppets: making representative shapes with the hands (mostly observed in silhouette) by one person, HAND

Foot Movement: deliberate foot movement not correspondent to moving across the screen or dancing, FEET

Hand circles: moving one or both hands in circles, not rolling the hands one over the other, HAND

Hand crossing: crossing the hands in front of the body, HAND

Hand Wave one hand: waving at the display only moving the hand (wrist movement) or only the lower arm, HAND

Hand Wave two hands: waving at the display only moving the hands (wrist movement) or only the lower arms, HAND

Hands In and Out: moving both hands closer and further apart from each other in front of the body, HAND

Hands Up and Down: bringing both hands up and down at the same time quickly; primarily happens as an elbow bend movement, HAND

Handshake/Holding hands: two people try to shake hands or hold their hands together HAND

High-five: two people try to high-five each other (success does not matter), HAND

Inviting: gesture used to induce someone else to engage; if it is noted, it is being conducted in a way to make sure it is on screen, ARM

Jump: single or multiple jumps in one place on the floor, WHOLE BODY

Jumping Jacks: jumping to have the legs spread and hands above the head, jumping back to have legs parallel and arms at side, WHOLE BODY

Kicking: kicking one or both feet in any direction, LEG

Leaning forward and back: tilting torso toward the screen and away from it; can be conducted with arms close to the body or held away from it, TORSO

Leaning left/right: tilting torso left and right; can be conducted with arms close to the body or held away from it, **TORSO**

Lunges: one leg is positioned forward with knee bent and foot flat on the ground while the other leg is positioned behind, **LEG**

Patty cake game: two people clapping their hands together rhythmically, **HAND**

Pivoting: twisting left and right while the feet stay in place, **TORSO**

Play fighting: any sort of gesture between two people meant to imitate fighting (may also be observed with inactive second avatar), **WHOLE BODY**

Posing: intentionally moving into a position and not intentionally moving out of it for a period of time (usually long enough to observe themselves), **WHOLE BODY**

Raising Both Arms: raising both arms above or around the shoulders and holding them there, **ARM**

Reach: stretching out one arm in a direction; different from wave because the arm does not move repeatedly or up and down, **ARM**

Rolling hands: repeatedly rotating the hands one over the other in front of the body, **HAND**

Shrugging: moving the shoulders up toward the ears and down, **SHOULDER**

Spinning: turning 360 degrees away from the display and to face it again, **WHOLE BODY**

Squatting: bending knees to move toward the floor, **WHOLE BODY**

Swipe: attempting to swipe hand to interact with the background/bringing hand quickly across the body, **ARM**

Swipe left to right/Swipe right to left: moving the hand quickly from one side to the other (usually across the midline of the body), **ARM/HAND**

Swipe up/down: moving the hand quickly upward or downward, **ARM/HAND**

Testing system limitations: moving in a way that addresses the edge cases of the representation and where it falls short (this will be different for all systems)

Turning: turning the entire body left and right (foot position must change), **WHOLE BODY**

Two-person shape: two participants try to create a shape together, **WHOLE BODY**

Walking side to side: walking from one spot on the floor to another, **WHOLE BODY**

Wide stance: standing with the feet/legs far apart, **WHOLE BODY**

References

1. 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; CHI '09; pp. 1083–1092. [\[CrossRef\]](#)
2. Lang, J.; Howell, E. *Researching UX: User Research*; SitePoint: Melbourne, Australia, 2017.
3. Salminen, J.; Jung, S.G.; Kamel, A.; Froneman, W.; Jansen, B.J. Who is in the sample? An analysis of real and surrogate users as participants in user study research in the information technology fields. *PeerJ Comput. Sci.* **2022**, *8*, e1136. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Ali, A.X.; Morris, M.R.; Wobbrock, J.O. Crowdlicit: A System for Conducting Distributed End-User Elicitation and Identification Studies. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019; CHI '19; pp. 1–12. [\[CrossRef\]](#)
5. Dourish, P. *Where the Action Is*; MIT Press: Cambridge, MA, USA, 2001.
6. Hornecker, E. The role of physicality in tangible and embodied interactions. *Interactions* **2011**, *18*, 19–23. [\[CrossRef\]](#)
7. Sorce, S.; Gentile, V.; Enea, C.; Gentile, A.; Malizia, A.; Milazzo, F. A touchless gestural system for extended information access within a campus. In Proceedings of the 2017 ACM SIGUCCS Annual Conference, Seattle, WA, USA, 1–4 October 2017; pp. 37–43.
8. Cafaro, F.; Roberts, J. Data through Movement: Designing Embodied Human-Data Interaction for Informal Learning. *Synth. Lect. Vis.* **2021**, *8*, 1–127.
9. Alt, F.; Geiger, S.; Höhl, W. ShapelineGuide: Teaching mid-air gestures for large interactive displays. In Proceedings of the 7th ACM International Symposium on Pervasive Displays, Munich, Germany, 6–8 June 2018; pp. 1–8.
10. Villarreal-Narvaez, S.; Vanderdonckt, J.; Vatavu, R.D.; Wobbrock, J.O. A Systematic Review of Gesture Elicitation Studies: What Can We Learn from 216 Studies? In Proceedings of the 2020 ACM Designing Interactive Systems Conference, Eindhoven, The Netherlands, 6–10 July 2020; DIS '20; pp. 855–872. [\[CrossRef\]](#)
11. Vogiatzidakis, P.; Koutsabasis, P. Gesture Elicitation Studies for Mid-Air Interaction: A Review. *Multimodal Technol. Interact.* **2018**, *2*, 65. [\[CrossRef\]](#)
12. Brignull, H.; Rogers, Y. Enticing people to interact with large public displays in public spaces. *Proc. Interact* **2003**, *3*, 17–24.

13. Müller, J.; Alt, F.; Michelis, D.; Schmidt, A. Requirements and design space for interactive public displays. In Proceedings of the 18th ACM International Conference on Multimedia, Firenze, Italy, 25–29 October 2010; pp. 1285–1294.
14. Jordan, B.; Henderson, A. Interaction analysis: Foundations and practice. *J. Learn. Sci.* **1995**, *4*, 39–103. [[CrossRef](#)]
15. Ali, A.X.; Morris, M.R.; Wobbrock, J.O. Distributed Interaction Design: Designing Human-Centered Interactions in a Time of Social Distancing. *Interactions* **2021**, *28*, 82–87. [[CrossRef](#)]
16. Pilourdault, J.; Amer-Yahia, S.; Roy, S.B.; Lee, D. Task relevance and diversity as worker motivation in crowdsourcing. In Proceedings of the 2018 IEEE 34th International Conference on Data Engineering (ICDE), Paris, France, 16–19 April 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 365–376.
17. Cafaro, F.; Panella, A.; Lyons, L.; Roberts, J.; Radinsky, J. I see you there! Developing identity-preserving embodied interaction for museum exhibits. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 27 April–2 May 2013; pp. 1911–1920.
18. Elmqvist, N. Embodied human-data interaction. In Proceedings of the ACM CHI 2011 Workshop Embodied Interaction: Theory and Practice in HCI, Vancouver, BC, Canada, 7–12 May 2011; pp. 104–107.
19. Cafaro, F. Using embodied allegories to design gesture suites for human-data interaction. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing—UbiComp '12, Pittsburgh, PA, USA, 5–8 September 2012; p. 560. [[CrossRef](#)]
20. Trajkova, M.; Alhakamy, A.; Cafaro, F.; Mallappa, R.; Kankara, S.R. Move Your Body: Engaging Museum Visitors with Human-Data Interaction. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–13.
21. Victorelli, E.Z.; Dos Reis, J.C.; Hornung, H.; Prado, A.B. Understanding human-data interaction: Literature review and recommendations for design. *Int. J. Hum. Comput. Stud.* **2019**, *134*, 13–32. [[CrossRef](#)]
22. Mortier, R.; Haddadi, H.; Henderson, T.; McAuley, D.; Crowcroft, J. Human-data interaction: The human face of the data-driven society. *arXiv* **2014**, arXiv:1412.6159.
23. 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; CHI EA '05; pp. 1869–1872. [[CrossRef](#)]
24. Morris, M.R. Web on the Wall: Insights from a Multimodal Interaction Elicitation Study. In Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces, Cambridge, MA, USA, 11–14 November 2012; ITS '12; pp. 95–104. [[CrossRef](#)]
25. Magrofuoco, N.; Vanderdonckt, J. Gelicit: A Cloud Platform for Distributed Gesture Elicitation Studies. *Proc. ACM Hum. Comput. Interact.* **2019**, *3*, 1–41. [[CrossRef](#)]
26. Paolacci, G.; Chandler, J.; Ipeirotis, P.G. Running experiments on Amazon Mechanical Turk. *Judgm. Decis. Mak.* **2010**, *5*, 411–419. [[CrossRef](#)]
27. McInnis, B.; Cosley, D.; Nam, C.; Leshed, G. Taking a HIT: Designing around Rejection, Mistrust, Risk, and Workers' Experiences in Amazon Mechanical Turk. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; CHI '16; pp. 2271–2282. [[CrossRef](#)]
28. Johnson, D.; Ryan, J.B. Amazon Mechanical Turk workers can provide consistent and economically meaningful data. *South. Econ. J.* **2020**, *87*, 369–385. [[CrossRef](#)]
29. Pittman, M.; Sheehan, K. Amazon's Mechanical Turk a digital sweatshop? Transparency and accountability in crowdsourced online research. *J. Media Ethics* **2016**, *31*, 260–262. [[CrossRef](#)]
30. Cafaro, F.; Lyons, L.; Antle, A.N. Framed guessability. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018.
31. Greenberg, S.; Boring, S.; Vermeulen, J.; Dostal, J. Dark patterns in proxemic interactions. In Proceedings of the 2014 Conference on Designing interactive systems, Vancouver, BC, Canada, 21–25 June 2014.
32. Rubegni, E.; Gentile, V.; Malizia, A.; Sorce, S.; Kargas, N. Child-display interaction: Exploring avatar-based touchless gestural interfaces. In Proceedings of the 8th ACM International Symposium on Pervasive Displays, Palermo, Italy, 12–14 June 2019; pp. 1–7.
33. Ackad, C.; Tomitsch, M.; Kay, J. Skeletons and Silhouettes. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016. [[CrossRef](#)]
34. Gentile, V.; Malizia, A.; Sorce, S.; Gentile, A. Designing touchless gestural interactions for public displays in-the-wild. In Proceedings of the Human-Computer Interaction: Interaction Technologies: 17th International Conference, HCI International 2015, Los Angeles, CA, USA, 2–7 August 2015; Proceedings, Part II 17; Springer: Berlin/Heidelberg, Germany, 2015; pp. 24–34.
35. Mishra, S.; Cafaro, F. Full body interaction beyond fun: Engaging museum visitors in human-data interaction. In Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction, Stockholm, Sweden, 18–21 March 2018; pp. 313–319.
36. Cheung, V.; Watson, D.; Vermeulen, J.; Hancock, M.; Scott, S. Overcoming interaction barriers in large public displays using personal devices. In Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces, Dresden, Germany, 16–19 November 2014; pp. 375–380.

37. Tomitsch, M.; Ackad, C.; Dawson, O.; Hespanhol, L.; Kay, J. Who Cares about the Content? An Analysis of Playful Behaviour at a Public Display. In Proceedings of the International Symposium on Pervasive Displays, Copenhagen, Denmark, 3–4 June 2014; PerDis '14; pp. 160–165. [\[CrossRef\]](#)
38. 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; CHI '12; pp. 297–306. [\[CrossRef\]](#)
39. Khamis, M.; Becker, C.; Bulling, A.; Alt, F. Which One is Me? Identifying Oneself on Public Displays. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; CHI '18; pp. 1–12. [\[CrossRef\]](#)
40. 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; UbiComp '15; pp. 1227–1238. [\[CrossRef\]](#)
41. Kittur, A.; Chi, E.H.; Suh, B. Crowdsourcing User Studies with Mechanical Turk. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy, 5–10 April 2008; CHI '08; pp. 453–456. [\[CrossRef\]](#)
42. Dutta, A.; Zisserman, A. The VIA Annotation Software for Images, Audio and Video. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019; MM '19. [\[CrossRef\]](#)
43. Ali, A.X.; Morris, M.R.; Wobbrock, J.O. Crowdsourcing similarity judgments for agreement analysis in end-user elicitation studies. In Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, Berlin, Germany, 14 October 2018; pp. 177–188.
44. Aşçı, S.; Rızvanoğlu, K. Left vs. right-handed UX: A comparative user study on a mobile application with left and right-handed users. In Proceedings of the Design, User Experience, and Usability. User Experience Design for Diverse Interaction Platforms and Environments: Third International Conference, DUXU 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, 22–27 June 2014; Proceedings, Part II 3; Springer: Berlin/Heidelberg, Germany, 2014; pp. 173–183.
45. Thimbleby, H. User interface design: Generative user engineering principles. In *Fundamentals of Human–Computer Interaction*; Elsevier: Amsterdam, The Netherlands, 1985; pp. 165–180.
46. Mackamul, E. Improving the Discoverability of Interactions in Interactive Systems. In Proceedings of the CHI Conference on Human Factors in Computing Systems Extended Abstracts, New Orleans, LA, USA, 29 April–5 May 2022; pp. 1–5.
47. Carter, S.; Mankoff, J.; Klemmer, S.R.; Matthews, T. Exiting the cleanroom: On ecological validity and ubiquitous computing. *Hum. Comput. Interact.* **2008**, *23*, 47–99. [\[CrossRef\]](#)
48. Dahlbäck, N.; Jönsson, A.; Ahrenberg, L. Wizard of Oz studies: Why and how. In Proceedings of the 1st International Conference on Intelligent User Interfaces, Orlando, FL, USA, 4–7 January 1993; pp. 193–200.
49. Sandifer, C. Time-based behaviors at an interactive science museum: Exploring the differences between weekday/weekend and family/nonfamily visitors. *Sci. Educ.* **1997**, *81*, 689–701. [\[CrossRef\]](#)
50. Ojala, T.; Kostakos, V.; Kukka, H.; Heikkinen, T.; Linden, T.; Jurmu, M.; Hosio, S.; Kruger, F.; Zanni, D. Multipurpose interactive public displays in the wild: Three years later. *Computer* **2012**, *45*, 42–49. [\[CrossRef\]](#)
51. Coenen, J.; Claes, S.; Moere, A.V. The concurrent use of touch and mid-air gestures or floor mat interaction on a public display. In Proceedings of the 6th ACM International Symposium on Pervasive Displays, Lugano, Switzerland, 7–9 June 2017; pp. 1–9.
52. Lakin, J.L.; Jefferis, V.E.; Cheng, C.M.; Chartrand, T.L. The Chameleon Effect as Social Glue: Evidence for the Evolutionary Significance of Nonconscious Mimicry. *J. Nonverbal Behav.* **2003**, *27*, 145–162. [\[CrossRef\]](#)
53. Morris, M.R.; Danieleescu, A.; Drucker, S.; Fisher, D.; Lee, B.; Schraefel, M.; Wobbrock, J.O. Reducing legacy bias in gesture elicitation studies. *Interactions* **2014**, *21*, 40–45. [\[CrossRef\]](#)
54. 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; TEI '16; pp. 86–91. [\[CrossRef\]](#)
55. Pérez-Mármol, J.M.; García-Ríos, M.C.; Ortega-Valdivieso, M.A.; Cano-Deltell, E.E.; Peralta-Ramírez, M.I.; Ickmans, K.; Aguilar-Ferrándiz, M.E. Effectiveness of a fine motor skills rehabilitation program on upper limb disability, manual dexterity, pinch strength, range of fingers motion, performance in activities of daily living, functional independency, and general self-efficacy in hand osteoarthritis: A randomized clinical trial. *J. Hand Ther.* **2017**, *30*, 262–273. [\[PubMed\]](#)
56. McNeill, D. *Language and Gesture*; Cambridge University Press: Cambridge, UK, 2000; Volume 2.
57. Norman, D.A. The way I see IT signifiers, not affordances. *Interactions* **2008**, *15*, 18–19. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.