Framed Guessability: Improving the Discoverability of Gestures and Body Movements for Full-Body Interaction

Francesco Cafaro
Human-Centered Computing
Indiana University-Purdue
University Indianapolis, USA
fcafaro@iu.edu

Leilah Lyons
Computer Science / Learning
Sciences, University of Illinois
at Chicago & New York Hall of
Science, USA
llyons@uic.edu

Alissa N. Antle
Interactive Arts and Technology
Simon Fraser University,
Vancouver, Canada
aantle@sfu.ca

ABSTRACT
The wide availability of body-sensing technologies (such as Nintendo Wii and Microsoft Kinect) has the potential to bring full-body interaction to the masses, but the design of hand gestures and body movements that can be easily discovered by the users of such systems is still a challenge. In this paper, we revise and evaluate Framed Guessability, a design methodology for crafting discoverable hand gestures and body movements that focuses participants' suggestions within a "frame," i.e. a scenario. We elicited gestures and body movements via the Guessability and the Framed Guessability methods, consulting 89 participants in-lab. We then conducted an in-situ quasi-experimental study with 138 museum visitors to compare the discoverability of gestures and body movements elicited with these two methods. We found that the Framed Guessability movements were more discoverable than those generated via traditional Guessability, even though in the museum there was no reference to the frame.

Author Keywords
Embodied Interaction; Full-Body Interaction; Human-Data Interaction; Frames; Elicitation; Guessability; Museums.

ACM Classification Keywords
H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

INTRODUCTION
"The human-machine interface is generalized beyond traditional control devices to permit physical participation with graphic images." [20] With these words, Myron Krueger described the first example of full-body interaction: Videoplace. Developed between 1977 and 1985, Krueger's Videoplace was an interactive installation that enabled two people (in different rooms) to interact together using hand gestures and body movements [19]. For many years, however, full-body interaction has been confined within the walls of research labs. In 1986, Buxton imagined a future archeologist digging up a personal computer and, because of its design, thinking that the user had one dominant hand, limited hearing, and no legs [5]. As England also noticed [11], the wide availability of body-sensing technologies might drastically change this scenario. Most recently, off-the-shelf sensors, such as Nintendo Wii and Microsoft Kinect, have the potential—through in-home applications and interactive installations in shared, public spaces—to bring full-body interaction to the masses.

ACM ISBN 978-1-4503-5620-6/18/04...$15.00.
https://doi.org/10.1145/3173574.3174167

The design of control actions (i.e., hand gestures and body movements) that can be easily *discovered* by to the users of full-body systems, however, remains problematic. As Norman observed, "a pure gestural system makes it difficult to *discover* the set of possibilities" that the system affords [27]. The users of full-body systems generally have no clue on the gesture or body movement that they can perform. This is particularly challenging for museums: visitors cannot consult user manuals before interacting with exhibits, and they often leave if the system is not responsive to the gestures and body movements that they attempt to do.

We believe that part of the *discoverability* problem that plagues full-body systems is rooted in the way in which we currently design hand gesture and body movements. Those control actions have been traditionally defined by designers, but it is hard to ensure that what designers have in mind is then *discoverable* at all to actual users. Guessability (or Elicitation) studies [35] are an alternative approach: in a lab setting, participants are exposed to an "effect" (something that the system can do), and then asked to recommend a gesture to control that effect. A winning suite is then constructed by selecting –for each effect –the control action that was recommended by the biggest number of people. Because guessability studies are conducted in-lab, there is no contextual information available to participants, and people are free to suggest completely unrelated control actions for different effects. This leads to disconnected sets of user-generated gestures and body movements. Thus, when the system is implemented in-situ, *discovering* a first control action does not necessarily provide any clue on what the others might be.

In this paper, we revise and evaluate Framed Guessability, a design methodology –initially introduced in [7] –that allows to design suites of *interconnected* hand gestures and body movements. Contrarily to traditional Guessability, Framed Guessability is based on the intuition of priming participants with a "context." In the revised version of Framed Guessability that we introduce in this work, we provide such "context" by priming participants to an elicitation study with a "frame" (i.e., a scenario). When they reason according to a frame, people know what may happen in that frame, and what cannot happen [21]: for example, one does not expect to board a plane in the middle of a highway. This approach allows to elicit hand gestures and body movements that are *interconnected*, because they are all consistent with one priming frame. Thus, we expected people to more easily progress from one control action to the others (in-situ), even when all the references to the priming frame used in-lab are removed. We want to highlight that the study in this paper is not an exploration of employing priming within final designs; rather, it evaluates a design technique, Framed Guessability that improves the "discoverability" of hand gestures and body movements by using priming within the design process.

In order to evaluate Framed Guessability, we first consulted 89 participants (in-lab), to elicit control actions via traditional Guessability and Framed Guessability for six effects. The frames we used were not strictly related with the application domain (data interaction). We then performed an in-situ study at the New York Hall of Science, with 138 museum visitors, to compare the *discoverability* of control actions (hand gestures and body movements) generated with those two approaches.

BACKGROUND & RELATED WORK

Frames

Frames were initially defined by Fillmore in the context of linguistics [12,13]. According to Fillmore, words in every languages have the power to evoke complex scenarios [12,13]. For instance, the word "restaurant" brings to mind being seated, seeing a waiter, being served food, etc. [17]. Fillmore's most recent work on frames focused on the creation of Framenet [3], an online database[1] that has been used for natural language processing (e.g., [4]) and semantic analysis (e.g., [1]). In [21,22], Lakoff incorporates Frames within the theory of Embodied Schemata [16]. According to Lakoff, frames are conceptual units that organize the basic experiences of our everyday life. They are actual brain circuits (i.e. neural connections) that are embodied in our brains [22].

Priming

Priming is an exposure to stimuli that evokes a frame, by activating a network of connected brain circuits [22]. Although a full analysis of existing priming methodologies is beyond the scope of this paper, we noticed two recurrent priming techniques in cognitive science.

Visual Priming: participants are exposed to pictorial representations of a scenario. For example, in [23], a group of students were prompted with either an image of an active journey (the picture of a man walking along a path) or of a passive journey (the picture of a moving train). Students primed with the active journey frame were more confident on which actions to take to achieve academic success.

Written Tasks: participants are asked to write something related with a scenario. For example, in [18], people were asked to recall five episodes from their past, and generate a keyword representing each episode; a first subset of participants was then provided with a picture of a path, and asked to put the keywords along that path, while a second subset was just asked to order the keywords in chronological order. In a follow-up questionnaire, people in the path condition showed an increased perception that what they did in their past shaped their present.

Traditional, Top-Down Approach to Gesture Design

As we mentioned in the Introduction, gestures are traditionally defined by designers, using sets of kinesthetic guidelines. For example, [33] describes a model for designing "grasp-sensitive objects." A "grasp" gesture is defined by five "meaningful factors" that should be considered when implementing a grasp-sensitive user interface: our Goal (e.g., the intended use of a screwdriver), our mental Relationship with the object (e.g, a tissue that belongs to us vs. a tissue used by another person), the Anatomy of our hand (e.g., palm size), the Setting where the object is located (e.g. a bottle on a shelf vs. in a case), and other Properties of the object (e.g, its size).

As observed in [36], a common risk of adopting a top-down approach is to produce guidelines that either don't suit user expectations, or are so rooted in a narrow context of use that they are ill-suited for generating sets of gestures for other contexts.

Guessability Studies: a Bottom-Up Approach

Wobbrock's Guessability studies [35] present an alternative approach: end-users, not experts, should be responsible for the design of interactive systems.

During a Guessability study, participants are exposed (one at a time) to a series of "effects" (i.e., functionalities of the system), and asked to recommend a "symbol" (or gesture) to produce that effect. Initially defined for the design of EdgeWrite [34] (a unistroke text-entry method for stylus-based tablets), Guessability studies have since then been applied to elicit control actions for a variety of contexts, including touch-screens [36], augmented reality [29], and wall-displays [24].

Participants to a Guessability study are free, however, to suggest completely unrelated gestures for different effects. An interesting example is illustrated in [31]: at the end of a

[1] https://framenet.icsi.berkeley.edu/fndrupal/

guessability study to elicit leap gestures for a TV, the user-generated gesture set included "moving hand to left" to go to the next channel, "drawing the letter M" to open a menu, and "waving one hand" to hide the menu.

Furthermore, the quality of the user-generated gesture set is measured at the end of the in-lab experiment using the "agreement" metric introduced in [35]. Even if the agreement metric is assumed to be a proxy for how intuitive control actions are, there is no experimental evidence that higher agreement (in-lab) results in gestures and body movements that are more *discoverable* (in-situ).

Framed Guessability: Providing a Context to Participants to an Elicitation Study

The revised version of Framed Guessability that we introduce in this paper is inspired by the methodology defined in [7]. Framed Guessability was introduced with the aim of increasing the "users' agreement" (in-lab) on the hand gestures and body movements that people recommend during an elicitation study [7]. Differently from traditional Guessability, participants to a Framed Guessability study are primed with a context. Specifically, the methodology introduced in [7] is structured in two phases. In Phase One (elicitation), a first group of participants took part in a traditional guessability study, with the purposed of eliciting control actions for 12 effects of a full-body system. After that, in Phase Two (priming & reduction), a second group of volunteers was exposed to a TV that was camouflaged to look like a mirror, and a live camera video feed from a Microsoft Kinect was displayed on the bottom half of the screen, to make the screen act as a "virtual mirror". This attempt to create a direct correspondence between physical space and virtual space [6] (to make them both look like a mirror) was called a "virtual mirror allegory" [7]. Participants to Phase Two were told to imagine that they were in front of a mirror and asked to choose their favorite gesture or body movement for each effect [7], among those that were recommended in Phase One by the first group of participants (i.e., participants to Phase Two were never asked to come up with their own recommendations).

The "old" Framed Guessability describe in [7], however, has three shortcoming that would compromise our effort to assess the impact of Framed Guessability in-situ, on the "discoverability" of control actions.

First, showing only a limited, pre-populated set of options to choose from (rather than leaving people free to recommend any gesture or body movement that they can imagine) automatically limits their choices. Thus, it is unclear if the "old" Framed Guessability [7] increases the users' agreement (on the actions that different people recommend) because of the priming, or just because participants have a reduced number of options to choose from.

Second, the procedure in [7] uses non-primed elicitation during Phase One, to generate control actions that are presented to participants in Phase Two. In this paper, we want to assess if priming participants with a context (i.e., a frame) during an elicitation study produces suites of more discoverable control action, because the context itself creates an invisible thread than connects the various control actions in the user-generated suite. The lack of a "context" during Phase One of the "old" Framed Guessability [7] (which is structured as a traditional guessability study, i.e. participants are not primed) may produce disconnected actions, which may in turn compromise our effort.

Third, the 1:1 relationship between the physical space (screen camouflaged as a "mirror") and the virtual space (camera-feed used as a virtual "mirror") might be difficult to reproduce in different scenarios -e.g., it would be hard to "dress up" the screen to make it look like a gym, and it may be distracting to add a gym representation on the screen near an unrelated data visualization.

FRAMED GUESSABILITY: PRIMING + ELICITATION

In the revised version of Framed Guessability that we introduce in this paper, we adopted three priming techniques to communicate a scenario since the beginning of the elicitation study. Specifically, the methodology that we used for our experiment is structured in two phases: Priming Phase, and Elicitation Phase. Both phases take place in a lab setting, with the same group of participants.

Volunteers are interviewed one at a time, according to the following procedure:

- *Priming (Phase one).* In the first phase of Framed Guessability, a participant is primed with a frame (e.g., "Eating a steak at a restaurant"). Priming is achieved with a combination of: (1) Visual Priming –pictures of the scenario are displayed on a screen as a slideshow; (2) Written Task –participants are asked to write things that they would do in the scenario; and, (3) Embodied Priming –people are asked to re-enact what they wrote in their written task.
- *Elicitation (Phase two).* The second phase of Framed Guessability immediately follows the priming phase, and is structured as a traditional Guessability study. The same participant is exposed to a sequence of effects (i.e., functionalities of the system), and ask to recommend one gesture and one body movement to "control" that effect.

As in traditional Guessability studies (e.g., [35]), the control action that has been recommended by the highest number of participants –for each effect –is included in the final suite of user-generated control actions.

PROBLEM STATEMENT: FRAMES & DISCOVERABILITY

The work that we present in this paper investigates the following research question: Is there a difference on the average number of control actions that users can *discover* when they interact (in-situ) with winning suites generated with Framed Guessability vs. the one created with Traditional Guessability?

HDI Task that the museum visitor can accomplish, and learning goal		Functionalities that the interactive system will provide	"Effect" used in the evaluation of Framed Guessability
Identify data that user controls	Increase sense of control on dataset (understand that the system is interactive)	Provide immediate, noticeable, and proportional response to movements	**Data Jiggle** –The data perform a subtle (but noticeable) oscillation.
Distill properties of single dataset	Compare single data elements to whole (reason about proportions)	Change aggregation scale	**Aggregation Level of Data** – Aggregate and disaggregate data elements by county, city, and census tracts
	See where lots of data elements are present (reason about quantities)	Emphasize areas of higher/lower density	**Data Up and Down** –Data move up, and they then fall down with a speed that is proportional to their size **Transparency Level of Data** – Allows to see data underneath others in areas with high density of data
Compare two datasets	Contrast the size of two or more datasets (reason about proportions)	Compare to whole	**Split Data** – Separate two datasets that have been previously aggregated
Expose properties of data representation	Compare how data are represented (reason about data representation)	Manipulate the scale of the representation	**Rescale Data** –Allows users to see how data would look like at different scales

Table 1 Effects for Human-Data-Interaction that were used for evaluating Framed Guessability

Our hypothesis was that gestures and body movements elicited (in-lab) with Framed Guessability are more discoverable (in-situ) than those generated with Traditional Guessability. During the design phase (in-lab), activating a specific frame should result in people recommending control actions that are *interconnected* (i.e., related with each other), because they are all grounded on that frame. Thus, during the actual use of the system (in-situ), people will be able to discover those control actions more easily than in the case of disconnected ones produced with traditional Guessability –even when all references to the priming frame are removed. We did not expect that the control actions would be activated in a sequence.

APPLICATION SCENARIO: HUMAN-DATA INTERACTION WITH *COCENSUS*

Human-Data Interaction (HDI) provides one of the most suitable contexts in which to evaluate Framed Guessability. HDI is a term that was firstly introduced in [6] and [2] to refer to interactive museum installations that allow multiple people to collaboratively explore large sets of data using hand gestures and body movements. Subsequently, the definition of HDI has been extended in [26], [14] and [93] to include the process of making big sets of data (e.g., data streams that generally undergo long data mining processes) accessible to a broader audience. This democratization process aims to move large datasets out of research labs (the domain of data scientists), into public spaces and informal learning settings, such as museums.

Designing control actions for Human-Data Interaction is particularly challenging, however. If we plan to use a top-down approach (designers-based), there are no well-established cultural metaphors that can be used to inform the design (such as visualizing a steering wheel when creating a car simulator). On the other hand, if we try to implement a bottom-up strategy (traditional Guessability study) the novelty of the interaction can be disorienting for participants to an elicitation study, as they cannot refer to any already-standardize gestures. Thus, Framed Guessability has the potential to make a massive impact on the design of HDI applications.

In particular, we used CoCensus [30] for our evaluation of Framed Guessability. CoCensus is a prototype interactive exhibit that allows museum visitors to explore a map-based visualization of US Census data, which is displayed on a big, shared screen (Figure 1).

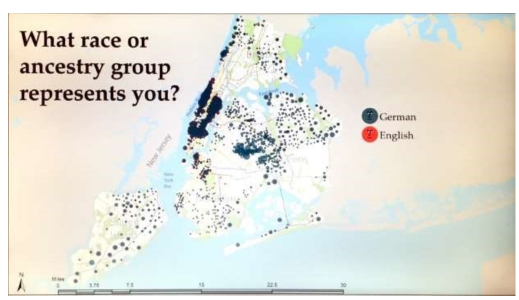

Figure 1. Visualization of US Census data in CoCensus. In this screen, people interact with scaled centroids representing German and English ancestry groups, on a map of New York.

Before entering the exhibit space, visitors are asked to fill out a reduced version of the US Census form, using a touch-screen kiosk. After approaching the interactive screen, each user is able to see her/his data, together with the data of other people in the room [30]. Using the ISYT tracking technology [9], each visitor is able to interact with her/his own slice of data, without interfering with the data of other participants.

For our evaluation of Framed Guessability, we focused on six effects described in Table 1: data jiggle, aggregation level of data, data up and down, transparency level of data, split data, and rescale data. They provide a first example of basic HDI tasks that museum visitors may want to accomplish when exploring large sets of data.

METHODOLOGY

The aim of Framed Guessability is to elicit (in-lab) suites of interconnected control actions that are *discoverable* (in-situ). First, we elicited suites of user-generated control actions using both traditional Guessability and Framed Guessability, for each of the six effects described in Table 1: data jiggle, aggregation level of data, data up and down, transparency level of data, split data, and rescale data. Because of the novelty of Framed Guessability, we decided to use two different priming frames (FUNHOUSE and GYM), in order to address the risk that we might have chosen a particular frame that is just more likely to lead to discoverability. As a frame selection strategy, we used frames that had this invariant: the user was the main actor (as our exhibit had them looking at a representation of their own data). Other application domains may use other invariants. This generated three experimental conditions: two Framed Guessability conditions (FUNHOUSE and GYM), and one Traditional Guessability condition (CONTROL). In-lab, we primed participants to the FUNHOUSE condition with the following frame: "*being in front of a distorted mirror while visiting a funhouse*." People in the GYM conditions were instead primed with: "*exercising at a gym with a weighted ball*." We expected to elicit a majority of full-body movements in the GYM condition, and a more balanced combination of hand gestures and body movements in the FUNHOUSE condition. Participants to the CONTROL condition were not primed, i.e. they skipped the priming phase and immediately started with the elicitation phase.

A total of 227 people participated to this study. For the in-lab sessions, we recruited N=89 undergraduate and graduate students using university mailing lists. Participants were compensated with a $5 gift card. For the in-situ experiment, we recruited N=138 museum visitors. Visitors were invited by a museum interpreter to approach our 90" screen and check out a new prototype exhibit.

In-Lab: Generating Suites of Control Actions with Framed Guessability and Traditional Guessability

In-lab, we divided N=89 volunteers into three experimental groups: n=29 for FUNHOUSE, n=31 for GYM, and n=29 for CONTROL.

Phase One: Priming

Figure 2, Figure 3, and Figure 4 illustrate how we primed participants to the GYM condition: (1) *visual priming*, i.e. a slideshow with evocative pictures of the GYM scenario were displayed on a screen when the participant entered the experimental room (Figure 2); (2) *written task*, i.e. the same participant was given a small worksheet and asked to write five things that she/he would do at a gym, while exercising with a weighted ball (Figure 3); and, (3) *embodied priming*, i.e. the participant was asked to re-enact the five things that she/he listed on her/his worksheet (Figure 4).

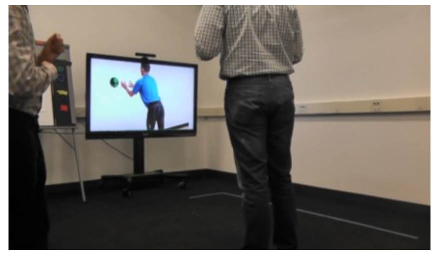

Figure 2. Experimental setup for the Priming Phase of Framed Guessability. A 65" display showed a slideshow with evocative pictures of the GYM frame.

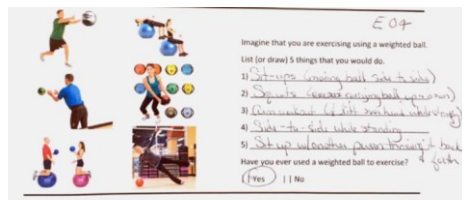

Figure 3. Worksheets used in the Priming Phase of Framed Guessability. Participants were asked to write five things that they would do in that scenario.

Figure 4. The last element of the Priming Phase of Framed Guessability is the Embodied Priming. In this picture, a participant re-enacts what she had listed on the worksheet for GYM in front of the screen.

Phase Two: Elicitation

The elicitation phase was structured as a traditional Guessability study. Regardless to the experimental condition (FUNHOUSE, GYM, or CONTROL), we showed the six "effects" from Table 1 to each participant, individually. We randomized the order in which effects were shown to each participant. After each effect, we asked the participant to recommend one gesture and one body movement to "control" that effect. At the end of the in-lab experiment, the control action that was recommended by the highest number of participants –in each experimental condition (i.e, FUNHOUSE, GYM, and CONTROL) –was selected to be part of the user-generated suite of control action for that experimental condition. Those suites are reported in the Results section.

In-Situ: Assessing the Impact of Framed Guessability on the Discoverability of Suites of Control Actions

The in-situ experiment was conducted at the New York Hall of Science with N=138 museum visitors: n=52 for FUNHOUSE, n=38 for GYM, and n=48 for CONTROL.

The focus of our evaluation was on the *discoverability* of control actions elicited (in-lab) using Framed Guessability; we did not want to just confirm that priming may have an impact when visiting a museum. Thus, museum visitors had no idea of what the frame used was: there were absolutely no cueing or priming in the exhibit portion of the work.

A museum interpreter invited visitors to try a new prototype exhibit using hand gestures and body movements. A 90” display showed the same visualization that we used in-lab.

A moderator near the exhibit asked museum visitor to “find out” what hand gestures and body movements they could use to control the visualization.

The in-situ study was structured to mirror the in-lab study: participants interacted with the system in either the FUNHOUSE, GYM, or CONTROL condition. Participants in the FUNHOUSE conditions were supposed to discover gestures and body movements that were generated in-lab by participants to the FUNHOUSE condition; likewise, people in the GYM and CONTROL conditions were expected to find the corresponding suites of control actions. The effects were only revealed to participants as soon as one participant was able to guess one of the control actions. As stated above, museum visitors did not receive any sort of priming; rather, they were all exposed to the same visualization on the screen.

Figure 5 shows two museum visitors interacting with the screen in the Funhouse condition, and properly guessing one of the hand gestures, moving one arm up and down, to control one of the effects (Data Up and Down).

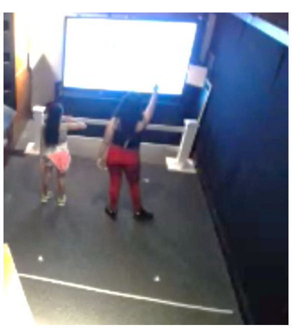

Figure 5 Two museum visitors interact with the prototype exhibit at the New York Hall of Science. A 90” display showed the same visualization that was used in the lab-study.

Because the focus of this study was on the *discoverability* of control actions –and not on the evaluation of tracking technologies –we used a Wizard-of-Oz approach to avoid the risk of the system missing or falsely recognizing hand gestures and body movements. A second moderator was hidden in a different room, and was able to see the interaction space using a webcam. Every time a visitor was able to “guess” a hand gesture or a body movement, the hidden moderator pressed a key on a wireless keyboard to trigger the proper effect.

RESULTS

We report the user-generated gesture suites of control actions (in-lab), describe the impact of Framed Guessability on their discoverability (in-situ), and discuss evidence of the lingering effect of the priming frame.

In-Lab: User-Generated Suites of Control Actions

User-Generated Suites of Control Actions

Table 2 reports the suites of user-generated control actions at the end of each experimental condition (FUNHOUSE, GYM, and CONTROL). The column labelled with *n* reports the percent of people who recommended that control action. Hand gestures are in green, body movements in blue.

As we expected, the proportion between hand gestures and body movements changed depending on the experimental condition. On a total of six control actions, three (50%) are body movements in the FUNHOUSE condition, five (83%) in the GYM condition, and two (33%) in the CONTROL condition. Future work should investigate this phenomenon, and evaluate which frames are best suited for eliciting body movements, and which ones are better for hand gestures.

Impact of the Frame on Generating Interconnected Suites of Control Actions (from Participants’ Remarks)

Participants’ remarks during the elicitation study reveal that the priming frame provided a common context not only for single effects, but across the entire elicitation experience. This is one example from the GYM condition:

> *User 61 –when describing a body movement for Transparency Level of Data. “You can also bring the ball down and then back up I would bring it down as it gets lighter and up as it gets darker.”*
>
> *Same user –when describing a hand gesture for Resize Data. “I can see bringing the ball out and forward I would bring it out to make it bigger and then in to get it smaller.”*
>
> *Same user –when describing a hand gesture for Aggregation Level of Data. “If you have the ball you can do two arms out and then they come together.”*

In other words, participants recommended gestures and body movements that were interconnected with each other, because they were grounded on the same priming frame. This resulted in different suites of user-generated control actions in FUNHOUSE and GYM.

EFFECT	FUNHOUSE		GYM		CONTROL	
	Control Action(s)	n	Control Action(s)	n	Control Action(s)	n
Data Up and Down	Move one hand up/down	48%	Jump	39%	Move one hand up/down	45%
Data Jiggle	Shake body	59%	Shake body	48%	Shake body	52%
Transparency	Walk backward/forward	31%	Walk backward/forward	26%	Move hand right/left	14%
Resize	Squat up and down	21%	Squat up and down	26%	Pinch and zoom gesture	28%
Aggregation Level	Arms apart/together (horizontally)	45%	Go up and down, make yourself smaller then extend body up	35%	Arms apart/together (horizontally)	41%
Split	Spread arms apart (horizontally)	79%	Spread arms apart (horizontally)	61%	Spread arms apart (horizontally)	45%

Table 2 Results of the In-Lab Study: Control Actions (Hand Gesture or Body Movement) that were recommended for each effect (by the highest number of participants), in the three experimental conditions (FUNHOUSE, GYM, CONTROL). "n" is the percent of participants that recommended that control action. Hand gestures are in blue, body movements in green.

In the absence of frames (i.e., CONTROL condition), people tended to imagine reference scenarios to recommend controlling actions for an effect. For example:

> *User 96 –Describing Data Up and Down. "Like aliens over New York. I would picture this like ET with the finger."*

Different participants, however, described different scenarios. Furthermore, same users generated multiple reference frames for the same interactive system. This resulted in disconnected control actions. For instance:

> *User 77 [for aggregation level of data]. "I imagine this game, I think it is like a board game …"*

> *User 77 [for split data]. I'm thinking of controlling each of them [the two datasets] with a handle if you turn left it would be coming closer or like this going apart."*

In-Situ: Effect on *Discoverability*

We report the number of control actions that were discovered in each experimental condition (FUNHOUSE, GYM, and CONTROL). Museum visitors were free to interact with the system: any reference to the original priming frame was removed during the in-situ experiment. The maximum number of control actions that participants could guess was six (i.e, the six effects in Table 1).

We analyzed the data from the in-situ study from two different perspectives. First, we considered all the experimental sessions that we recorded, which included people interacting as solo users, in groups of two people, and in groups of three people –we did not want to alter how people usually interact with other exhibits at that museum. Second, we investigated the number of control actions that were discovered by each user (rather than by all visitors within that experimental session). For uniformity, we considered only visitors interacting in pairs ("dyads" sessions), because the presence or absence of companions might alter discoverability

Control Actions Discovered in all Experimental Sessions
One-hundred and thirty-eight museum visitors took part in the experiment, for a total of 74 experimental sessions: 27 for FUNHOUSE, 22 for GYM, 25 for CONTROL. Figure 6 illustrates the percent of experimental sessions in which the cumulative number of control actions discovered by all participants was either zero, one, two, three, four, five, or six. The percent of experimental sessions in which participants were able to identify four and five control actions was 4% and 0% respectively in the traditional Guessability conditions. These values increased to 19% and 11% (FUNHOUSE), and 14% and 9% (GYM).

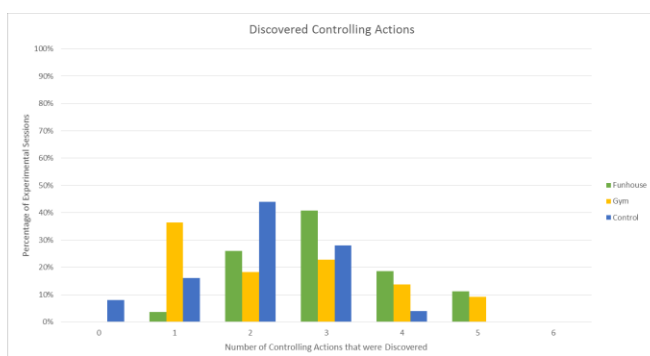

Figure 6 Percent of experimental sessions (on a total of 74 sessions) in which the cumulative number of control actions discovered by all participants is either one, two, three, four, five, or six.

A one-way ANOVA was conducted with the experimental condition (FUNHOUSE, GYM, and CONTROL) as independent variable, and the number of control actions that were discovered by all participants during an experimental session as dependent variable. There was a statistically significant difference between experimental conditions, as determined by ANOVA ($F(2,71) = 7.16$, $p<0.005$). A Tukey post-hoc test revealed that the number of control actions that were discovered was significantly higher in the FUNHOUSE ($M=3.07$, $SD=1.04$) than in the CONTROL condition ($M=2.04$, $SD=0.98$, $p<0.004$).

Control Actions Discovered by Each Participant

Frames and schemata inform the reasoning of each person. The ultimate goal of Framed Guessability is to improve the discoverability of gestures and body movements for *each* user: we want to provide all museum visitors with an opportunity to interact with the full-body system and with each other, without pushing anybody into the periphery of the social space. Thus, the third step of the discoverability analysis considered the number of control actions that were discovered by each participant.

For uniformity (we suspect that the social context can affect discoverability), we limited this analysis to the experimental sessions in which two visitors interacted together with the system ("dyads" sessions). Because we did not want to interfere with the natural way in which people usually interact with other exhibits at that museum, the number of sessions with a solo user and with groups of two people was not balanced across experimental conditions. When we consider only "dyads" sessions, we are left with a total of 52 experimental sessions: 17 for FUNHOUSE, 12 for GYM, 23 for CONTROL (104 total museum visitors). Thus, a total of N=104 total museum visitors participated in these "dyads" sessions: n=34 for FUNHOUSE, n=24 for GYM, n=46 for CONTROL.

A one-way ANOVA was conducted with the experimental condition (FUNHOUSE, GYM, and CONTROL) as independent variable, and the number of control actions that were discovered by each participant during an experimental session as dependent variable. There was a statistically significant difference between experimental conditions, as determined by ANOVA ($F(2,101) = 7.97$, $p<0.003$). A Tukey post-hoc test revealed that: (1) the number of control actions that were discovered by each participant was significantly higher in the FUNHOUSE ($M=2.59$, $SD=1.05$) than in the CONTROL condition ($M=1.74$, $SD=1.02$, $p<0.003$); and, (2) the number of control actions that were discovered by each participant was significantly higher in the GYM ($M=2.42$, $SD=1.38$) than in the CONTROL condition ($M=1.74$, $SD=1.02$, $p<0.04$).

Thus, museum visitors (individually) discovered a bigger number of control actions (hand gestures and body movements) in the two Framed Guessability conditions (FUNHOUSE and GYM), when compared to the traditional Guessability condition (CONTROL) -see Figure 7.

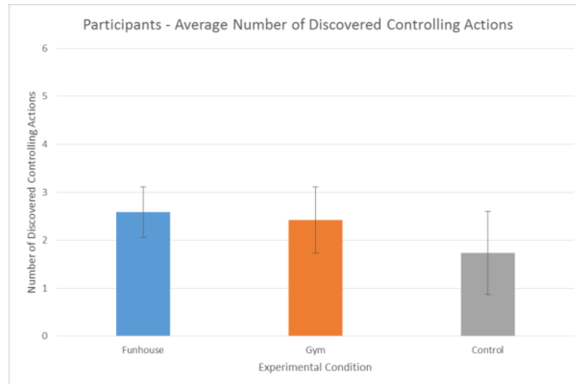

Figure 7 Average number of control actions that were discovered by each participant (in dyads sessions), under the three experimental conditions (FUNHOUSE, GYM, and CONTROL).

In-Lab + In-Situ: Evidence of the Lingering Effect of the Frame

During the in-lab study we noticed that, when different frames were used, participants were more or less likely to use hand gestures versus body movements. In particular, the GYM framing encouraged people to suggest more body movements than gestures. We looked for a similar pattern in the in situ use to see if there is a "residue" of the now-absent priming frame. Because we removed any reference to the frame, there should be no reason to expect different proportions of hand and body movements in the in-situ trials across framing conditions, unless there is a "spreading activation effect" of the schema as they discover how to control the system. In this section, we show how this is exactly what happened during the in-situ evaluation.

If we had a perfect understanding of the full network of schemata that are activated by a frame, we could look for those schemata, and patterns in how one schema might co-activate another. However, this is much more of a foundational question about how schema work, and would require a separate research project on embodied cognition.

Figure 8 shows the average number of hand gestures and body movements that each participant recommended in-lab for the six effects, in each of the three experimental conditions (FUNHOUSE, GYM, and CONTROL). Participants could make multiple suggestions for the same effect, and not to make any recommendation (skipping to the next effect) if they had no suggestion after having been prompted twice by the moderator.

A series of multivariate ANOVA (MANOVA) was conducted with the experimental condition (FUNHOUSE, GYM, and CONTROL) as independent variable, and the number of Hand Gestures and Body Movements that each participant recommended during the in-lab study as dependent variable. In the case of Hand Gesture, there was a statistically significant difference across experimental conditions, as determined by ANOVA ($F(2,83) = 14.89$, $p<0.0001$). A Tukey post-hoc test revealed that the number of hand gestures that each participant suggested was significantly higher in the CONTROL (M=5.70, SD=1.14, p<0.0001) and FUNHOUSE (M=4.76, SD=1.27, p<0.024) than in the GYM condition (M=3.80, SD=1.49). In the case of body movements, there was a statistically significant difference across experimental conditions, as determined by ANOVA (F(2,83) = 13.59, p<0.0001). A Tukey post-hoc test revealed that the number of Body Movements that each participant suggested was significantly higher in the Gym (M=5.10, SD=1.88) than in the Funhouse (M=3.69, SD=1.79, p<0.008) and Control (M=2.67, SD=1.62, p<0.0001) conditions.

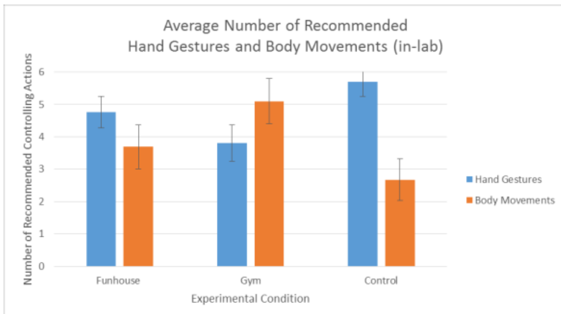

Figure 8. In-Lab. Average number Hand Gestures and Body movements that participants recommended in-lab, in each experimental condition (Funhouse, Gym, and Control).

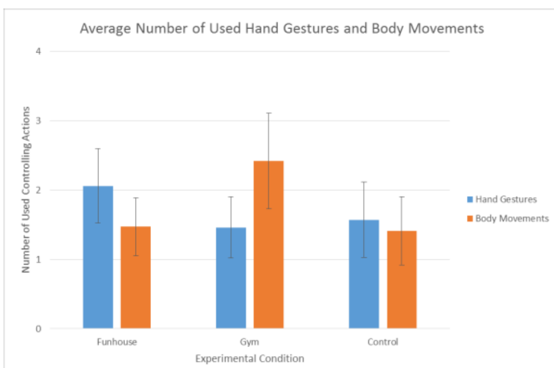

Figure 9. In-Situ. Average number of Hand Gestures and Body movements that participants attempted to use, in each experimental condition (Funhouse, Gym, and Control) - regardless of whether that action could trigger an effect.

Similarly, Figure 9 shows the average number of Hand Gestures and Body Movements that participants attempted to use –regardless of whether that action could trigger an effect on the screen. In this phase of the analysis, we did not want to re-assess discoverability; rather, we wanted to understand if there was a lingering effect of the frame (which may explain the increased discoverability of control actions when using Framed Guessability): this is the reason why we considered all gestures and body movements that participants attempted to do, rather than only those in the winning suite of control actions. A series of multivariate ANOVA (MANOVA) was conducted with the experimental condition (Funhouse, Gym, and Control) as independent variable, and the number of hand gestures and Body Movements that each participant "attempted to use" during an experimental session as dependent variable. In the case of hand gesture, there was a statistically significant difference across experimental conditions, as determined by ANOVA (F(2,101) = 3.31, p<0.05). In the case of body movements, there was a statistically significant difference across experimental conditions, as determined by ANOVA (F(2,101) = 8.88, p<0.001). A Tukey post-hoc test revealed that the number of body movements that each participant tried to use was significantly higher in the Gym (M=2.42, SD=1.38) than in the Funhouse (M=1.47, SD=0.83, p<0.003) and Control (M=1.41, SD=0.98, p<0.001) conditions. These results (from the in-situ evaluation) resemble the pattern on the number of hand gestures and body movements that participants recommended (in-lab) during the Framed Guessability study.

DISCUSSION

Framed Guessability and Embodied Schemata

When we designed Framed Guessability, we hypothesized that the frame-designed suites would outperform the CONTROL condition suite because even if visitors were not privy to the priming frame, the suites of recommended gestures were likely to be interconnected. By interconnected we mean that there would be a certain unifying logic to them, such that discovering some of the gestures would help visitors discover other gestures. This hypothesis rests on Lakoff's and Johnson's theory of embodied schemata [16][22], which suggests that if one schema is activated in the mind of a person, its activation can spread to other schema in an associative pattern formed via real-world experiences.

Aligning Frames with the Intended Context of Use

In this study, we selected the GYM and FUNHOUSE scenarios because they both involved people scenarios in which people move their own bodies, just as we wanted them to "move" the data visualized in the final exhibit; one biases more towards full-body movement (GYM) and the other more towards arm gestures and poses (FUNHOUSE). Our results show that frames influence participants' recommendation during an elicitation studies. Thus, we want to highlight three design recommendations that may be helpful to designers when selecting a frame: (1) Frames should preserve the relationship between physical movement and visual attention desired in the final design. In this work, we wanted visitors to carefully attend to how each physical movement altered an element visualized on-screen – which happens when preening or flexing in front of mirrors; whereas another design - e.g., a game where users move through a 3D space – might want users' visual focus on an on-screen goal while the body movements "blindly" support navigation, thus requiring more proprioceptive actions; (2) Frames should match the desired "continuity" of actions (whether actions should be discrete triggers or continuous input) [8]; and, (3) The frame should align with the social norms for body movement in the space where the system will be deployed (one reason the GYM scenario was less successful then FUNHOUSE may have been that GYM actions are not typically engaged in at museums, which could inhibit discovery).

Enactment Design [10] may provide insights on how the actions elicited during Framed Guessability are "situated" in the context (i.e., the frame) provided to participants. In particular, the User Enactments approach [28] provides a framework for creating "scenarios" that participants enact to engage in discussion concerning technological interventions. This approach may contribute to establishing guidelines for crafting frames for FG and should be incorporated in future work.

Hand Gestures, Body Movements, and Legacy Biases

As we expected, at the end of the GYM condition, the large majority (5 out of 6) of winning control actions were body movements, while the FUNHOUSE condition elicited more hand gestures (3 out of 6). In-situ, the novelty of full-body interaction might have had an impact on the discoverability of control actions, and favored the more balanced suite elicited with FUNHOUSE (which included a mix of hand gestures and body movements) over the one generated with GYM (mostly full-body movements). This should be investigated in future work.

The CONTROL condition included mostly hand gestures. We believe this is due to the impact of legacy biases [25]: 33% of people in the CONTROL condition made explicit reference to touch-screen devices, and at least four of the hand gestures in the winning suite of control actions were based on variations of the "swipe" and "pinch-and-zoom" gestures that are commonly implemented in tablets and phones. For the most part, participants to the two Framed Guessability conditions (FUNHOUSE and GYM) did not refer to their prior experience with technology when describing hand gestures and body movements. We observed an exception only for 7% of participants.

One might think that the legacy biases may provide an entry point to the interaction, because they lead participants to an elicitation study to recommend (in-lab) already standardized gestures that people will be familiar with (in-situ). Surprisingly, in our study, the control actions generated from legacy biases (in-lab) were not that discoverable (in-situ). For example, only 2 people (over 48 participants) were able to discover the "pinch and zoom" gesture in-situ (CONTROL). We believe that legacy biases did not work as well as Framed Guessability actions because: (1) they may be based on different technologies (e.g., people are not used to pinch and zooming in the air, while they frequently do so on mobile phones); and, (2) they were not connected with each other.

Embodied Ideation techniques like camouflaging the space, sound clues, and objects [15] could be used to reinforce priming in Framed Guessability. In particular, "Enactments strategies" are similar to the embodied priming in FG, and there are parallels between the "disruptive idea" of material "re-contextualization" [32], and the way in which frames re-contextualize the interaction (away from previously standardized gestures -i.e., legacy biases -e.g. pinch and zoom) in Framed Guessability. This should be incorporated in future work.

CONCLUSION AND FUTURE WORK

In this paper, we revised Framed Guessability and provided an evaluation of this novel methodology for designing hand gestures and body movements for full-body interaction. The revised version of Framed Guessability that we introduced in this paper incorporates "frames" into the elicitation process, in order to provide a unifying context for applications –such as Human-Data Interaction –in which there are not well-established design patterns or metaphors. During an elicitation study (in-lab), frames clue people in to which actions are more (or less) appropriate. The results of our experimental evaluation (in-situ –when all references to the original priming frame are removed) show that Framed Guessability significantly increases the *discoverability* of control actions, when compared to the gestures and body movements elicited with traditional Guessability. The value of this design methodology rests on the unusual result that higher discoverability was attained with no in-situ priming.

This work contributes to interaction design in two ways. (1) We illustrate that one facet that matters to the users of full-body systems (in terms of "discoverability") is establishing a connecting thread across control actions. This is in deep contrast with the idea that only standardized actions (e.g., pinch and zoom) can be easily discovered, and opens new opportunities for creatively crafting rich interaction designs. We will clarify this in the Conclusion. (2) Although we evaluated Framed Guessability with an interactive data visualization exhibit, it can be applied to a plethora of full-body interaction scenarios -e.g., interactive installations for cultural heritage, interactive public art, and commercial devices such as smart TVs, in which users do not typically consult manuals before interacting.

Future work should explore what happens when users are also primed in-situ, i.e. the degree to which the number and nature of cues incorporated into final design space (e.g., visual cues, or environmental design cues) can improve the discoverability of control actions. We also believe that Framed Guessability can be used in scenarios that go beyond human-computer interaction: for example, to catalog the schemata that are triggered by specific frames, and to understand interconnected reasoning patterns that may be triggered by gestures and body movements.

ACKNOWLEDGMENTS

A portion of this work was supported by the National Science Foundation under NSF INSPIRE 1248052: Studying and Promoting Quantitative and Spatial Reasoning with Complex Visual Data across School, Museum, and WebMedia Contexts. We want to thank Rebecca Eydt for her help with the data collection at NySci.

REFERENCES

1. James F Allen, Mary Swift, and Will De Beaumont. 2008. Deep semantic analysis of text. *Proceedings of the 2008 Conference on Semantics in Text Processing*, Association for Computational Linguistics, 343–354.
2. Alissa N Antle, Paul Marshall, and Elise van den Hoven. 2011. Embodied Interaction: Theory and Practice in HCI. *Proceedings of CHI 2011*.
3. Collin F Baker, Charles J Fillmore, and John B Lowe. 1998. The berkeley framenet project. *Proceedings of the 17th international conference on Computational linguistics-Volume 1*, Association for Computational Linguistics, 86–90.
4. Aljoscha Burchardt, Marco Pennacchiotti, Stefan Thater, and Manfred Pinkal. 2009. Assessing the impact of frame semantics on textual entailment. *Natural Language Engineering* 15, 4: 527–550.
5. W Buxton and B Myers. 1986. A Study in Two-Handed Input. *CHI 86 Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press, 321–326.
6. Francesco Cafaro. 2012. Using embodied allegories to design gesture suites for human-data interaction. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ACM, 560–563.
7. Francesco Cafaro, Leilah Lyons, Raymond Kang, Josh Radinsky, Jessica Roberts, and Kristen Vogt. 2013. Framed Guessability: Using Embodied Allegories to Increase User Agreement on Gesture Sets. *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction*, ACM, 197–204.
8. Francesco Cafaro, Leilah Lyons, Jessica Roberts, and Josh Radinsky. 2014. The Uncanny Valley of Embodied Interaction Design. *Proceedings of the 2014 conference on Designing interactive systems*, ACM, 1075–1078.
9. Francesco Cafaro, Alessandro Panella, Leilah Lyons, Jessica Roberts, and Josh Radinsky. 2013. I see you there!: developing identity-preserving embodied interaction for museum exhibits. *Proceedings of the 2013 ACM annual conference on Human factors in computing systems*, ACM, 1911–1920.
10. Scott Davidoff, Min Lee, Anind Dey, and John Zimmerman. 2007. Rapidly exploring application design through speed dating. *UbiComp 2007: Ubiquitous Computing*: 429–446.
11. David England. 2011. Whole Body Interaction: An Introduction. In D. England, ed., *Whole Body Interaction An Introduction*. Springer London, 1–5.
12. Charles Fillmore. 1982. Frame semantics. *Linguistics in the morning calm*: 111–137.
13. Charles J Fillmore. 1977. The case for case reopened. *Syntax and semantics* 8, 1977: 59–82.
14. Heiko Hornung, Roberto Pereira, M Cecilia C Baranauskas, and Kecheng Liu. 2015. Challenges for Human-Data Interaction–A Semiotic Perspective. In *Human-Computer Interaction: Design and Evaluation*. Springer, 37–48.
15. Philemonne Jaasma, Dorothé Smit, Jelle van Dijk, Thomas Latcham, Ambra Trotto, and Caroline Hummels. 2017. The Blue Studio: Designing an Interactive Environment for Embodied Multi-Stakeholder Ideation Processes. *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction*, ACM, 1–10.
16. Mark Johnson. 1987. *The Body in the Mind*. The University of Chicago Press.
17. Hank Johnston. 2002. Verification and proof in frame and discourse analysis. *Methods of social movement research*: 62–91.
18. Lucas A. Keefer, Mark J. Landau, Daniel Sullivan, and Zachary K. Rothschild. 2011. Exploring metaphor's epistemic function: Uncertainty moderates metaphor-consistent priming effects on social perceptions. *Journal of Experimental Social Psychology* 47, 3: 657–660.
19. Myron W Krueger. 1977. Responsive environments. *Proceedings of the June 13-16, 1977, national computer conference*, ACM, 423–433.
20. Myron W Krueger, Thomas Gionfriddo, and Katrin Hinrichsen. 1985. VIDEOPLACE An Artificial Reality. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 35–40.
21. George Lakoff. 2006. Simple framing. *Rockridge Institute* 14.
22. George Lakoff. 2008. The Neural Theory of Metaphor. *The Cambridge Handbook of Metaphor and Thought*: 17–38.
23. Mark J Landau, Daphna Oyserman, Lucas A Keefer, and George C Smith. 2014. The college journey and academic engagement: How metaphor use enhances identity-based motivation. *Journal of Personality and Social Psychology* 106, 5: 679–698.
24. Meredith Ringel Morris. 2012. Web on the wall: insights from a multimodal interaction elicitation study. *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces*, ACM, 95–104.
25. Meredith Ringel Morris, Andreea Danielescu, Steven Drucker, Danyel Fisher, Bongshin Lee, and Jacob O Wobbrock. 2014. Reducing legacy bias in gesture elicitation studies. *interactions* 21, 3: 40–45.
26. Richard Mortier, Hamed Haddadi, Tristan Henderson, Derek McAuley, and Jon Crowcroft. 2014. Human-data interaction: The human face of the data-driven society. *Available at SSRN 2508051*.

27. Donald A Norman. 2010. Natural User Interfaces Are Not Natural. *Interactions* 17, 3: 6–10.
28. William Odom, John Zimmerman, Scott Davidoff, Jodi Forlizzi, Anind K Dey, and Min Kyung Lee. 2012. A Fieldwork of the Future with User Enactments. *Proceedings of the Designing Interactive Systems Conference*, ACM, 338–347.
29. Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn. 2013. User-defined gestures for augmented reality. *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, ACM, 955–960.
30. Jessica A. Roberts, Leilah Lyons, Joshua Radinsky, and Francesco Cafaro. 2012. Connecting Visitors to Exhibits through Design: Exploring United States census data with CoCensus. *ICLS'12*.
31. Radu-Daniel Vatavu and Ionut-Alexandru Zaiti. 2014. Leap gestures for TV: insights from an elicitation study. *Proceedings of the 2014 ACM international conference on Interactive experiences for TV and online video*, ACM, 131–138.
32. Danielle Wilde, Anna Vallgårda, and Oscar Tomico. 2017. Embodied Design Ideation Methods: Analysing the Power of Estrangement. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ACM, 5158–5170.
33. Raphael Wimmer. 2011. Grasp sensing for human-computer interaction. *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, ACM, 221–228.
34. Jacob O Wobbrock. 2006. A robust design for accessible text entry. *SIGACCESS Access. Comput.* 84: 48–51.
35. Jacob O Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A Myers. 2005. Maximizing the guessability of symbolic input. *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, ACM, 1869–1872.
36. Jacob O Wobbrock, Meredith Ringel Morris, and Andrew D Wilson. 2009. User-defined gestures for surface computing. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 1083–1092.