Framed Guessability: Using Embodied Allegories to Increase User Agreement on Gesture Sets

Francesco Cafaro[1], Leilah Lyons[1,2], Raymond Kang[2], Josh Radinsky[2], Jessica Roberts[2], Kristen Vogt[2]
[1]Computer Science [2]Learning Sciences
University of Illinois at Chicago
{fcafar2, llyons, rkang, joshuar, jrober31, kvogt4} @uic.edu

ABSTRACT
Despite the wide availability of body-sensing technologies, the design of control gestures that feel natural and that can be intuitively "guessed" by the users is still an embodied interaction challenge. This is especially true for systems that require a *set* of complementary control gestures. Part of the problem lies in the separation between the *locus* of the interaction (the body) and the *focus* of the interaction (the screen). We extend Johnson's theory of Embodied Schemata with Embodied Allegories, in order to create a unifying context that spans across the locus and focus of interaction. We present results that demonstrate how this approach increases the chance that users select the same gesture or movement for producing an effect within the virtual context, and that the resultant gesture set is deemed more intuitive by users. We also present the accompanying methodology, "Framed Guessability," which can increase users' agreement when conducting Guessability Studies.

Author Keywords
Embodied Interaction, Embodied Schemata, Embodied Allegories, Guessability Study, User-Defined Gestures.

ACM Classification Keywords
D.2.2 Design Tools and Techniques: User Interfaces. H.1.2 User/Machine Systems: Human Factors. H.5.2 User Interfaces: Interaction styles (e.g., commands, menus, forms, direct manipulation), Theory and methods.

General Terms
Design, Human Factors, Theory.

INTRODUCTION
In the late 1970s, Myron Krueger introduced one of the first whole-body, interactive systems: Videoplace [12]. Participants' physical movements were used to "*navigate a computer defined visual space*." According to Krueger, what really matters for the design of such an interactive system is establishing a meaningful relationship between the user's actions and the system's response [12]. Forty years later, the wide availability of body sensing technologies, such as Nintendo Wii and Microsoft Kinect, opens up Whole-Body Interaction [8] to the masses. However, this means that more attention must be given to the Embodied [7] nature of these interactions. How to design control gestures that feel natural and that can be intuitively "guessed" by the users is still a challenge.

TEI'14, February 16 - 19 2014, Munich, Germany

ACM 978-1-4503-2635-3/14/02…$15.00.
http://dx.doi.org/10.1145/2540930.2540944

Top-Down, Designer-Determined Gestures
Gestures are usually defined by designers, with sets of domain-specific guidelines. For instance, [21] introduces a framework for "grasps," and [14] provides precise definitions of "tapping", "pressing" and "dragging," and shows how these three basic movements can be used to construct a variety of hand gestures. As observed in [23], however, this top-down approach runs the risk of producing guidelines that 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.

Bottom-Up, User-Derived Gestures
"Guessability Studies" [22] follow an alternative approach: end-users are exposed to an "effect" (something that the system is able to do), and are asked to recommend a gesture (or "symbol") for controlling that "effect". The "winning gestures" are those recommended by the biggest number of users –design guidelines and theory are not involved. Despite their contribution to democratizing the design process, the very definition of "winning gestures" illustatres how different people may suggest different symbols for the same effect.

Design Metaphors: Top-Down Meets Bottom-Up
In the early days of WIMP (Windows, Icon, Mouse, Pointer) systems, proponents of Direct Manipulation [18] recommended selecting an appropriate metaphor to guide the interaction design [19]: e.g., a "desktop", or tools like drawing programs' "paint brush." Metaphors have also been used in embodied systems, to create successful mappings between a well-known domain and the virtual world. For instance, [13] describes a projected screen resembling a car windshield for "navigating a graphic world." Users mimic the movements they would use with a car wheel, in order to set the direction in which they want to go. Because the design capitalizes on users' internalized, culturally-mediated

"scripts" for how to operate this kind of interface, this technique can be learned in few seconds [13]. Metaphors also give generative fuel to designers, who can "riff" off of already-established metaphors to add new functionality to systems (e.g., adding a "paint bucket").

Embodied Metaphors for Embodied Interaction

However, for those embodied systems which do not have a culturally-recognizable input "prop" like a steering wheel, the problem of metaphor selection is nontrivial: in addition to being culturally relevant, the metaphor must span and unite both the user's proprioceptive experience of the motions used to control the system as well as the on-screen response of the system. One promising approach to developing principled, "intuitive" embodied interaction has been to use the theory of *embodied schemata* [11]. Designers rely on kinesthetic metaphors developed by humans at a very young age through constructivist explorations of the physical world. For instance, [3] explores how an embodied schema can be used to teach an abstract concept by establishing a metaphorical connection between physical "balance" and "social justice". This approach works very well where there is a 1:1 correspondence between a specific concept and a specific body movement, but it doesn't yield the same type of generative, complementary design possibilities as culturally-defined metaphors (like the "desktop") can.

Contribution: Embodied Allegories for Gesture Suites

Part of the problem with designing for embodied interaction lies in the separation between the *locus* of the interaction (the body) and the *focus* of the interaction (the screen). The user needs to understand how to map the locus to the focus, a mapping that not only cues them as to which bodily actions might possibly be relevant to the digital context, but also helps them anticipate how those actions might affect the digital context. This kind of context-to-context parallelism is a key component of allegories, which are extended metaphors in which multiple symbolic objects, people, and concepts serve as "stand-ins" for real-world entities. Building on Cafaro's previous work [6], this paper expands the range of interactions that can be designed into a system via Embodied Schemata with the concept of Embodied Allegories: interactive, embodied experiences in which a suite of gestures is based on multiple embodied schemata. In an Embodied Allegory, there is a unifying, parallel "story" present in both the physical space and in the virtual world; this shared, common "context" (i.e. the allegory) can help to frame and constrain users' conceptions about what they might try to enact in the physical space and what those enactments might produce in the virtual realm.

We present the results of a study that demonstrates that Embodied Allegories can be used to frame embodied interaction experiences, in order to increase users' agreement on how to produce effects within the virtual context with their gestures and body movements. We also present the methodology we used, "Framed Guessability," as an approach that can offer improved utility for conducting guessability studies of sets of actions.

RELATED WORK

Here we provide more detail on Embodied Schemata, their implications for Human-Computer Interaction, other theories of gesture, and Guessability Studies.

Embodied Schemata and their Metaphorical Projection

In [11], Johnson defines an Embodied Schema as a recurrent pattern, shape and regularity in our daily body experience. Embodied schemata may be metaphorically projected, creating common structures to guide (and constrain) our understanding and reasoning. For example, "*balance*" is a bodily experience that a baby learns when she/he stands, falls to the floor, and stands again, until she/he understands how to distribute his or her mass (body weight) on supports (legs) to keep a balanced erect posture. This schema rapidly evolves metaphorically, so that we can unconsciously apply the concept of "balance" to different target domains, such as visual arts and architecture [11].

Embodied Schemata in Human-Computer Interaction

The theory of Embodied Schemata has been applied to the design of tangible user interfaces, such as [10], and to embodied systems, in order to create a connection between user's actions and abstract concepts. In Springboard [3], the concept of "balance" is explored in the domain of social justice, using a whole-body interactive environment. Sound Maker ([1,2]) connects the speed of body movements to sound to teach abstract music concepts such as "tempo."

However, it may be difficult to generalize the results of these works. In [1,2,3,4] there is a very linear 1:1 relation between gesture and effect. For instance, concepts like fast/slow tempo are linearly mapped to fast/slow movements. What happens when the learning value rests in the whole interactive experience, rather than on a single, atomic concept? The tight input-output coupling might mean that the metaphor breaks down or conflicts with other metaphors when one tries to expand the experience to incorporate a larger set of interactions.

Other Theories of Gesture

Apart from Embodied Schemata, there is a dearth of theory-grounded approaches that have been used to understand gestures. McNeill [15], for instance, classifies gestures into four distinct categories: *Iconics*, which are pictorial representations of a concept; *Metaphoric*, which represent an abstract concept; *Deictis*, which indicate objects being spoken about; *Beats*, to emphasize elements of a conversation. Embodied-interaction gestures might be considered as "*Emblems*" in the Kendon's Contninuum of human gestures: they need to be learned, and they are meaningful without speech, although they may also occur with speech [15]. As observed in [23], such frameworks were designed to analyze human-human discourse, not to *design* human-computer discourse. It is unclear the extent to which these framings can be used to design human-computer interaction gestures from first-principles.

An Atheoretical Approach: Guessability Studies

Guessability studies [22] are based on the idea that end-users, not experts, should be responsible for the design of gestures for interactive systems. Users are exposed to an "effect" (or "referent"), and then asked to freely recommend a "cause" (or "symbol"). For instance, [23] studied defining gestures for surface computing. Users were exposed to "referents" such as "selecting a shape" and "moving a shape" on the surface screen; they proposed "symbols" (i.e. gestures) such as "tap" and "drag".

The Agreement A is computed as a "*degree of consensus*" score among participants [22]:

$$A = \frac{\sum_{r \in R} \sum_{P_i \in P_r} \left(\frac{|P_i|}{|P_r|}\right)^2}{|R|}$$

where r is a referent in the set of all referents R, P_r is the set of proposed gestures for referent r, and P_i is a sub-set of almost identical gestures from P_r. For a single referent r, $A_r = \sum_{P_i \in P_r} \left(\frac{|P_i|}{|P_r|}\right)^2$. For example, if we interview 7 participants and 2 people select "clap" while 5 choose "wave hand" to trigger a given effect, P_r={"clap"(x2), "wave hand"(x5)}, and $A_r=(2/7)^2+(5/7)^2=0.59$. Although originally defined for designing symbol sets for PDA handwriting recognition [22], this measure of agreement has been used in a range of domains, from mobile phones [17], to augmented reality [16].

It is worth noting that guessability studies are based on two assumptions: (1) Users' behavior is rarely "systematic enough" to follow the same principles and categories that designers use to classify gestures [23]; (2) McNeill's categories cannot be used to inform the design of gestures for interactive systems, because they are based on human discourse [23]. If one subscribes to these assumptions, gesture design cannot be grounded a-priori on a specific theory; on the contrary, designers should rely on a pragmatic, statistical analysis of user's preferences.

In this paper, we challenge this theory-agnostic approach, in the context of embodied interaction: (1) We show how the gestures that people recommend during a guessability study are often *related* –i.e. multiple gestures can be viewed as deriving from common Embodied Schemata; (2) We use Embodied Allegories to constrain users' reasoning and increase the guessability of the user-defined gesture set S.

EMBODIED ALLEGORIES

In this section, we define "Embodied Allegories" by first describing "allegories." (It is worth noting that we use the definition of "allegory" that positions it as an extended metaphor. Many people conflate allegories with aphoristic fables, but the definition we use has no "moral" to the story. Rather, "allegories" are extended metaphors which act as "unifying suites" for multiple metaphors). We then review some characteristics of embodied schemata that we use for defining Embodied Allegories, and their implications for the design of gestures and body movements.

Plato's Cave: Allegories as Unifying Suites

The allegory of Plato's Cave was presented in "*The Republic"* to illustrate the dualistic distinction between ideas and the material world perceived through our senses. Some *captives* (a metaphor for "mankind") have been kept in chain inside a *cave* (a metaphor for the "material world") since they were born. Behind them, a fire illuminates puppets –and all the captives can see are the *shadows* of those puppets (a metaphor for the "how we experience objects through our senses") on the wall in front of them. Other metaphors are present which are not detailed here.

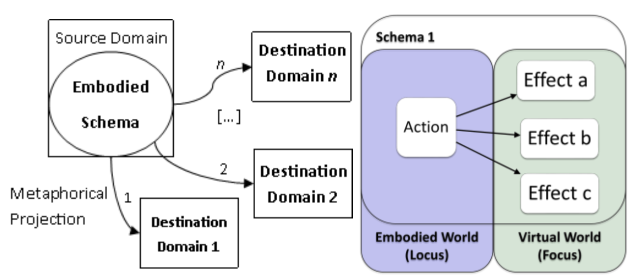
An allegory can work as a "container" for multiple, coherent metaphors – metaphors such as the sun (i.e., the philosophy), the shadows (i.e., our experience of objects), etc., acquire a precise, specific meaning because the reader sees them within a well-defined allegorical framework (i.e., the story of the cave). Using this allegory, readers can extend it to cover other metaphors: for example, the *puppeteers* who select and parade objects before the fire may be seen as a metaphor for politicians or other people who wish to manipulate the thoughts of others.

Embodied Schemata: Embodiment, Polysemy, and Constraint on Reasoning

"Embodied Allegories" extend the theory of Embodied Schemata. Before being able to provide a definition of "Embodied Allegories," we need to highlight some characteristics of Johnson's theory [11]:

Embodiment. The embodied experience generates a metaphorical projection involving bodily orientations, movements, perceptions, and manipulation of objects.

Polysemy (support for Multiple Metaphorical Projections). Each Embodied Schema may generate multiple metaphorical projections (Figure 1.a). These metaphors are all constrained by the schema and share a multiple *related* meanings. For instance, the Embodied Schema "PATH" can be metaphorically projected into "Purposes are Physical Goals" (e.g., being on your way of getting a PhD), but also inform our mathematical reasoning (e.g., a number line).



(a) **(b)**

Figure 1. (a) One Embodied Schema may generate multiple Metaphorical Projections into different (but related) Destination Domains. (b) When the source domain is the physical space, and the destination domain is the virtual world, the polysemy of embodied schemata may contribute in creating different expectations for users.

We believe that the polysemy of embodied schemata is one of the reasons why the users of embodied systems may have difficulties in mapping their actions in the physical space (locus of interaction) into an effect in the virtual space (focus of interaction) – see Figure 1.b.

Constraint on Reasoning. Embodied Schemata provide structures that constrain our *understanding* and *reasoning*. In particular, different metaphors and schemata trigger different inference models. For example, people may use either a "water-flow model" vs. a "crow model" when reasoning about electricity; this leads to different approaches to hydraulic problems.

***Embodied Allegories*: A Framing Context Spanning the Locus and Focus of Interaction**

We define Embodied Allegories as *interactive, embodied experiences* in which an allegory (i.e. an extended metaphor) works as unifying frame for a set of embodied schemata. The allegory can be used to provide a framing context which spans across both the physical space and the virtual world. Gestures and body movements are metaphorical projections of the embodied schemata that are contained within this allegorical framing. Defining characteristics for the use of Embodied Allegories are:

Embodiment spans Locus and Focus. An interactive, embodied experience designed with an Embodied Allegory involves not only gestures and body movements in the physical space (locus of interaction), but also the visualization, sounds, and other system feedback in the virtual space (focus of interaction). These two spaces, the embodied-world locus of interaction and the virtual-world focus of interaction, comprise two parallel worlds which need to be mapped on to one another. Embodied schemata underpin this mapping process.

Heterosemetic Constraint on Schema Mapping. While a given Schema may be polysemetic, if the framing context induces heterosemy (by reducing the possibilities to only one sensible mapping) we can constrain the space of likely interpretations. We believe that Embodied Allegories may serve as this framing context. Thus, some embodied schema mappings that may be triggered during the interaction with a generic version of an embodied system no longer seem relevant when an Embodied Allegory is introduced.

Complementary Constraint on Suite Membership. Activating a particular schema within an Embodied Allegory inhibits the activation of schemata which might provide a conflicting mapping between actions and effects. This should limit the suite of schemata to a complementary, non-overlapping set. Thus, if a suite of gestures was selected via an Embodied Allegory, the complementarity of the gestures within the suite should be recognizable even if the Embodied Allegory is never made explicit.

HYPOTHESES

Our hypotheses were the following:

(H1) Embodied schema theory suggests that a given schema can map to multiple related domains (polysemy). Traditional guessability studies ask users to map a target domain (the domain of possible virtual effects) to a source domain (the domain of possible gestures). We posit that, although the guessability methodology is itself atheoretical, users implicitly draw on embodied schemata when generating possible gestures for an effect. We further claim that this would be evidenced by the fact that the sometimes-large results of a regular guessability study can be classified using a smaller set of embodied schemata. More specifically, the k gestures that users recommend during a guessability study are the metaphorical projection of e embodied schemata (Figure 2), and so $k < e$.

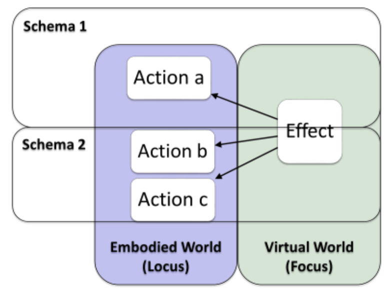


Figure 2. During a guessability study, users are exposed to an element of the destination domain (an effect), and try to reconstruct the source domain (an action in the physical world). Here, Actions b and c are variations of the same underlying schema (Schema 2).

(H2) The use of an Embodied Allegory can constrain the metaphorical reasoning that people do when interacting with an embodied system (i.e., moving away from a polysemetic mapping towards a heterosemetic mapping). Thus, when Embodied Allegories are introduced, the users of an embodied system will base their reasoning only on a subset q of the e embodied schemata on which the gestures recommended during the guessability study were based. This increases the agreement score, because it limits the number of gestures that people recommend during a guessability study, by providing a unifying context across the physical space and the virtual realm (Figure 3).

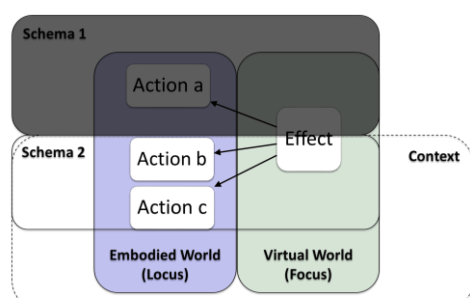


Figure 3. Embodied Allegories constraint the metaphorical understanding of the interactive experience; some embodied schemata do not "make sense" anymore within the allegorical framework. This reduces the number of recommended gestures and thus increases the agreement score.

(H3) An allegory contains complementary metaphors. We posit that once a suite of gestures is defined using an Embodied Allegory, users should be able to recognize the complementarity of the suite even without being told about the unifying allegory (i.e., without being "primed").

METHODS

The study took place in three separate sessions, to investigate the three hypotheses. We recruited university students as participants: 8 in the first session, 8 in the second, and 13 in the third, for a total of 29 people. All the interviews were videotaped, and two moderators were always present. Only one user at a time was allowed in the room to avoid influencing other participants. Nobody who participated in the earlier sessions was recruited for subsequent sessions.

Use Scenario: Whole-Body Interaction with Census Data

The testbed for this study is a prototype museum exhibit (*CoCensus*) designed to allow visitors to explore visualization of data (e.g., US Census data) via embodied interaction [5]. When entering the interaction space, a user selects an ancestry category to control (e.g., German). Approaching a 65" LCD screen, she/he sees "her/his" data subset (e.g., the number of German immigrants, represented as scaled "bubbles") on a shared display, along with the data subsets of the other visitors in the room (Figure 4). To allow visitors to productively explore the data we need a set of gestures and body movements that can be "intuitively" used to control the visualization of visitors' data subsets.

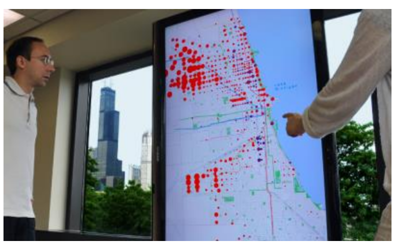


Figure 4. Two users interact with *CoCensus*. The map shows scaled centroids ("bubbles") representing the population of each person's self-identified ethnic group.

We did not disclose our goals with the participants to our study (as recommended in [23] for guessability studies). We simply told them that they were exploring the size of different ethnic groups in a large Midwestern city.

Phase One: Guessability Study

The first phase of the experiment was structured as a Guessability Study [22]: users were first exposed to an effect (or "referent"), and then asked to recommend a cause (or "symbol"). The purpose of Phase One was to verify if participants in a guessability study recommend gestures that are *related*, i.e. based on common embodied schemata (H1).

The screen was set-up to show a map of the city's area, and bubble data that represented the size of the Indian and German population. A moderator escorted a participant into the room, and briefly explained what was shown on the screen. Then, the moderator showed 12 different animations of possible dynamic manipulations of the data visualization (e.g., moving the bubbles, fading them in and out, rescaling them). We randomized the order in which the animations were shown. After each animation, we asked the participant which gesture or body movement she/he would like to do to trigger or to control that effect. We asked participants to think aloud, and to enact the gesture or movement that they recommended, to try them out. We used field notes, videos of the gesture enactment, and audio recordings of participants' rationales to categorize each proposed gesture by its likely underlying embodied schema.

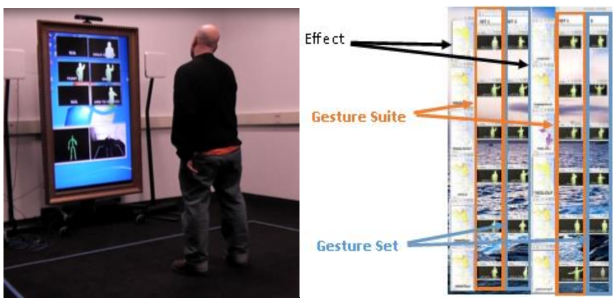
Phase Two: Assessing the Framing Effect of Embodied Allegories

Phase Two was designed in order to evaluate the effect of Embodied Allegories on users' preferences for the gestures that people generated during Phase One. After we supplied an allegorical framing, we expected users to disagree with a subset of metaphorical projections that people made during phase one; in other words, we expected participants of phase two to *systematically* dislike those gestures/body movements generated during phase one which were not in keeping with the allegory. We say *systematically* because, for each effect, we expected the least favorite gestures to be *related*, in that they are metaphorical projections of a common embodied schema. Our hypothesis – see (H2) – was that some embodied schemata no longer "make sense" when an allegorical framing is used. This is why we dub this stage of the methodology "framed guessability."

For the second phase of our study, we selected a "virtual mirror" as our allegory, because the Census is often seen as an "imperfect mirror" of our society. We used three strategies to emphasize this allegory to participants, to prime them to view possible interaction mappings through this framing lens: (1) *Visualization*. We decided to include a silhouette of the user and the camera's RGB output, so that the participant was able to see her/his own "reflection" on the screen; (2) *Verbal instruction*. The moderator described the screen as a virtual mirror; (3) *Physical design*. The display was mounted vertically within a gilded wooden frame to more closely resemble a mirror.

At the beginning of the interview, one moderator clearly stated that the screen was supposed to be a "mirror (see Figure 5.a) and pointed out the live silhouette and camera views, inviting the user to wave her/his hand and to see "what it looks like."

Participants were then asked to select which of the candidate gestures would most (and least) suit each target effect (an animation of how the data visualization would respond). To do so, we used a two-step process of gesture rehearsal and gesture selection. For the gesture rehearsal for each effect, we presented all gestures proposed for that effect by the participants in phase one. Each candidate gesture was demonstrated with a movie clip of an actor showing the gesture in silhouette (see Figure 5.a). For each animation, we asked participants to enact each gesture/body movement three times, or until they felt familiar with it, using the live feed of their own motion at the bottom of the screen to help them match their performance to that of the actor in the movie clip. We randomized the window in which each candidate gesture movie clip was shown (in case proximity of the clip to the live feed "reflection" of the user would affect their choices). After rehearsal, we displayed a full-screen version of the target effect (an animation), and asked the user which gesture/body movement was her/his "favorite", and which was her/his "least favorite", to control or trigger that effect. We repeated the effect (i.e. the animation) at least twice, and we then switched back to the rehearsal screen, so that participants were able to review gestures/body movements and see what they "looked like" in the camera view. Users were asked to think aloud, and to explain the reason of their choices.



(a) (b)

Figure 5. (a) During Phase Two, participants were exposed to all the candidate gestures per target effect (here, 6 windows at top of the screen). They were able to see a live camera feed and their silhouette (bottom of the screen) as they rehearsed each candidate gesture at least 3 times. (b) During Phase Three, the gestures of the guessability set and of the allegorically-framed guessability suite were shown side by side, together with the effect they were supposed to control/produce.

Phase Three: Confirmation of Gesture Suite

The gestures most preferred by participants in the allegorically-framed Phase Two were consolidated into a suite. To confirm that the gestures in this suite are complementary, we performed a comparative study of this suite against the set of gestures that would have been selected by the traditional guessability study of Phase One. We followed the guidelines presented in [23] in order to construct the "winning gesture set" from Phase One: for each effect, we chose the most recommended (or one of the most recommended) gesture; in case of conflicts (same winning gesture for two different effects), we assigned the gesture to the effect where it had the highest score. This resulted in a different but valid gesture set to compare against the suite derived from Phase Two. It should be noted that the set and the suite partially overlapped (a few highly-popular gesture mappings were the same).

During this phase, we did not mention the mirror allegory at all, and we removed the user's silhouette and the RGB camera view from the visualization. On the 65" LCD screen, we showed all the 12 data visualization effects (animations) in two columns. There were two columns next to each animation: one with the gestures from the gesture suite (allegorically-framed guessability study), the other with the gestures from the gesture set (traditional guessability study) –see Figure 5.b.

We told participants that we needed to decide between the first and the second set of gestures; we asked them which set they "preferred" and found "more intuitive" as a whole.

RESULTS AND DISCUSSION

In this section, we report the results of the three phases of our experiment. We then present a case study of a target effect (data jiggle), to better explain our methodology.

H1: User-Defined Gestures are related by common Embodied Schemata

The n=8 people that we interviewed during Phase One proposed an average of k=4.54 gestures/body movements per effect (SD=1.01). A researcher who was deeply familiar with embodied schemata coded each gesture with the most likely underlying embodied schemata (drawing from those listed in [11]). For each effect, we were able to identify an average of e=2.15 embodied schemata (SD=0.77). (Given the small number of applicable schemata we deemed it unnecessary to train another researcher to perform this coding, but for richer interaction spaces we would recommend doing so and checking for inter-coder reliability). As per hypothesis H1, we observed that, for each effect, $k>e$: for each effect, subsets of suggested gestures/body movements were *related* by a metaphorical projection of a common embodied schema (see Table 1).

H2: Embodied Allegories may increase the Guessability of the Users-Defined Gesture Set

At the end of Phase One, the agreement of the winning gesture set (i.e. those included in the set S) was A_1=0.32 (SD=0.09). After introducing the mirror allegory (Phase Two), the agreement increased to A_2=0.52 (SD=0.20). Note that users in Phase Two were not generating gestures based on the allegory: they were selecting from amongst the gestures generated "blind" in Phase One. As per hypothesis (H2), the agreement on the gestures of the allegorically-framed gesture suite was significantly higher than the agreement on the gesture set of the traditional guessability study (t(22)=3.248, p<0.0037, one tailed).

H3: A Gesture Suite Developed with an Embodied Allegory will be Cohesive and Complementary

When comparing the gesture set derived from the traditional guessability study in Phase One against the gesture suite derived from the allegorically-framed guessability study of Phase Two, 70% of people (9 of 13) expressed a preference for the gesture suite. In other words, the suite of gestures assembled via the framed guessability study seems to be more intuitive than the set constructed from the traditional guessability study, even though both were assembled from the same pool of gestures. Note that this pool of gestures was generated in the absence of an allegory, and that in Phase Three none of the participants were made aware of the allegory used to assemble the suite. The participants' preference for the allegorical suite shows that they were able to detect the underlying allegorical cohesion of the suite even when uncued as to its nature.

Case Study: Data Jiggle

In this animation, the bubbles were quickly oscillating right and left, introducing a sort of "jiggle" in the data (an effect we wished to use in the final exhibit to allow users to know which data set they were controlling). Table 1 reports the gestures that were proposed by users during Phase One (guessability study), together with the Embodied Schema and metaphors that we used to classify them.

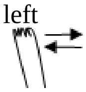
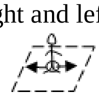
Enactment	Metaphor	E. Schema
Stop sign 	*Hand gesture* –The bubbles are free to jiggle until I block them	FORCE (blockage)
Move arm right and left 	*Hand movement* –The bubbles follow the path I draw with my hands	PATH (right/left)
Walk right/left 	*Body movement* –The bubbles walk with me (they follow my path in the space)	
Move torso right and left 	*Body movements* –The bubbles dance with me (they follow the path I draw with my torso)	

Table 1. Classification of gestures that were proposed to control the "data jiggle" during the guessability study. The embodied schema were taken from [11].

During the guessability study, people interpreted the effect (the jiggle) and the consequence of their action (gesture/body movement) in two different ways. A first group of users considered the bubbles as independent objects, with their own "bouncing" force. They interpreted their task as stopping and restarting the jiggle. For instance, one person said: "*I want it [the jiggle] to be done. I want to stop it.*" The second group of users, instead, believed that the bubbles were following their own gestures/body movements: they "*emulate my motion in the physical space.*" In other words, both the user and the bubbles were following the same PATH, right and left. The user had control on the starting and end point of the path, the trajectory, the speed, and the intermediate points. In Table 2, we use Gentner's *analogical model* [9] to better explain the metaphorical projection from the user's schemata into the virtual space.

The use of two different embodied schemata produces very different mental models, and distinct inference patterns. Users who reasoned according to the FORCE (Blockage) schema interpreted their gesture as a "on/off" switch; those who used PATH introduced a direct mapping between their position in the space and the x,y coordinates of the bubble.

Schema: FORCE (Blockage) *Metaphor*. The bubbles freely move (jiggle) on the screen; I can stop that movement with my hand. *Object mapping*. Bubbles->Floating objects; My hand->On/Off switch *Property mapping*. Jiggle (On/Off) -> Stop sign(On/Off) *Relationship imported*. On/Off->On/Off
Schema: PATH (Right/Left) *Metaphor*. The location of the bubbles on the screen is a physical location (of my hand, body, or torso). *Object mapping*. Bubbles->My hand, or body, or torso *Property mapping*. x,y (screen) -> x,y,z (physical space) *Relationship imported*. Motion along a path (screen) -> Motion along a path (physical space)

Table 2. Metaphorical projection from the user action (source domain) to the virtual space (destination domain).

Table 3 compares the results of Phase 1's guessability study with the scores of each gesture/body movement after we introduced the mirror Embodied Allegory (Phase 2).

Enactment	**Phase 1**	**Phase 2: Best**	**Phase 2: Worst**
Stop sign	3	0	9
Move arm	3	2	0
Walk	2	1	0
Move torso	3	6	0

Table 3. Number of users who proposed a gesture/body movement in Phase One, vs. number of users who considered it as their favorite ("best") or least favorite ("worst").

At the end of Phase One, three gestures/movements ("stop sign", "move arm", "move torso") all achieved the same score. After the mirror allegory introduction in Phase Two, however, participants unanimously judged "stop sign" as the worst gesture for producing the effect. It seems the mirror allegory inhibited the selection of the FORCE (Blockage) schema from the users' reasoning and inference. This is not surprising, as a person expects to see her/his own reflection when she/he is front of a mirror; and other reflected objects do not have a life on their own. At the end of Phase Two, agreement for the most popular selection ("Move torso") increased from 0.22 to 0.51 (+132%).

CONCLUSION AND FUTURE WORK

In this paper, we extended Johnson's theory of Embodied Schemata [11] with Embodied Allegories. We showed how Embodied Allegories can be used to constrain users' reasoning during the interaction with an embodied system, and we proved how they become a pragmatic tool for increasing the agreement in a user-defined gesture set. We also demonstrated that the suite of gestures created via our allegorically-framed guessability methodology are more likely to be judged coherent and complementary than a set assembled via a traditional guessability study, even when the users have no idea about the underlying allegory.

We believe that the use of Embodied Allegories supports a stronger and more intuitive relationship between the physical input (locus of interaction) and the system's output (focus); this is, in a sense, an approach to the "Platonic ideal" of Direct Manipulation, to the extent that "*the user is able to apply intellect directly to the task; the tool itself seems to disappear*" [18]. To demonstrate this, further studies should investigate if the effect of the increased agreement is inversely correlated with the time and effort required to progress through the levels of interaction described in [20]: from the initial understanding that the system is interactive, to the ultimate goal of being able to instrumentally operate it.

ACKNOWLEDGMENTS

This material is based upon work 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 Web-Media Contexts.

REFERENCES

1. Antle, A.N., Corness, G., Bakker, S., Droumeva, M., Van Den Hoven, E., and Bevans, A. Designing to support reasoned imagination through embodied metaphor. *Proceeding of the seventh ACM conference on Creativity and cognition CC 09*, (2009), 275.
2. Antle, A.N., Droumeva, M., and Corness, G. Playing with The Sound Maker : Do Embodied Metaphors Help Children Learn ? *Proceedings of the 7th international conference on Interaction design and children IDC 08*, (2008), 178–185.
3. Antle, A.N. Springboard : Exploring Embodiment , Balance and Social Justice. *Social Justice A Journal Of Crime Conflict And World Order*, (2009), 3961–3966.
4. Bakker, S., Antle, A.N., and Van Den Hoven, E. Embodied metaphors in tangible interaction design. *Personal Ubiquitous Comput. 16*, 4 (2012), 433–449.
5. Cafaro, F., Panella, A., Lyons, L., Roberts, J., and Radinsky, J. 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 (2013), 1911–1920.
6. Cafaro, F. Using embodied allegories to design gesture suites for human-data interaction. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ACM (2012), 560–563.
7. Dourish, P. *Where the action is: the foundations of embodied interaction*. MIT Press, Cambridge, MA, USA, 2001.
8. England, D. Whole Body Interaction: An Introduction. In D. England, ed., *Whole Body Interaction An Introduction*. Springer London, 2011, 1–5.
9. Gentner, D. Structure-mapping: A theoretical framework for analogy. *Cognitive Science 7*, 2 (1983), 155–170.
10. Hurtienne, J. and Israel, J.H. Image schemas and their metaphorical extensions. *Tangible and Embedded Interaction*, ACM Press (2007), 127.
11. Johnson, M. *The Body in the Mind*. The University of Chicago Press, 1987.
12. Krueger, M.W. Responsive environments. *Proceedings of the June 13-16, 1977 National Computer Conference*, ACM (1977), 423–433.
13. Krueger, M.W. Environmental technology: making the real world virtual. *Commun. ACM 36*, 7 (1993), 36–37.
14. Lao, S., Heng, X., Zhang, G., Ling, Y., and Wang, P. A gestural interaction design model for multi-touch displays. *Proceedings of the 23rd British HCI Group*, British Computer Society (2009), 440–446.
15. McNeill, D. *Hand and Mind: What Gestures Reveal about Thought*. University Of Chicago Press, 1992.
16. Piumsomboon, T., Clark, A., Billinghurst, M., and Cockburn, A. User-defined gestures for augmented reality. *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, ACM (2013), 955–960.
17. Ruiz, J., Li, Y., and Lank, E. User-defined motion gestures for mobile interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2011), 197–206.
18. Rutkowski, C. An Introduction to the Human Applications Standard Computer Interface, Part 1: Theory and Principles. *Byte 7*, 11 (1982), 291 –310.
19. Shneiderman, B. Direct Manipulation: A Step Beyond Programming Languages. *Computer 16*, 8 (1983).
20. Williams, A., Kabisch, E., and Dourish, P. From interaction to participation: Configuring space through embodied interaction. *Proceedings of Ubicomp 2005, LNCS 3660*, (2005), 287–304.
21. Wimmer, R. Grasp sensing for human-computer interaction. *Proceedings of TEI '11*, ACM (2011), 221–228.
22. Wobbrock, J.O., Aung, H.H., Rothrock, B., and Myers, B.A. Maximizing the guessability of symbolic input. *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, ACM (2005), 1869–1872.
23. Wobbrock, J.O., Morris, M.R., and Wilson, A.D. User-defined gestures for surface computing. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2009), 1083–1092.