Advance Access publication on 12 February 2015
doi:10.1093/iwc/iwv003

# What is Intuitive Interaction? Balancing Users' Performance and Satisfaction with Natural User Interfaces

ANNA MACARANAS, ALISSA N. ANTLE* AND BERNHARD E. RIECKE

*Department of Interactive Arts and Technology, Simon Fraser University, Surrey, British Columbia, Canada*
**Corresponding author: aantle@sfu.ca*

**Designers of natural user interfaces are faced with several challenges when creating interaction models for controlling applications, including the wide range of possible input actions and the lack of affordances, which they can use to design controls. In order to contribute to the development of design guidelines in this design space, we conducted an exploratory, mixed methods study. We investigated three top-down approaches to designing intuitive interaction mappings for a whole body system implemented with camera vision. These were metaphoric, isomorphic and 'everyday' or conventional. In order to identify some of the benefits and limitations of each approach, we compared the designs based on measures of usability, intuitiveness and engagement with the material represented in the system. From our study, we found that while the metaphoric design enhanced users' performance at completing tasks, the lack of discoverability of the interaction model left them feeling incompetent and dissatisfied. We found that the isomorphic design enabled users to focus on tasks rather than learning how to use the system. Conversely, designs based on previous conventions had to be learned, had a time cost for the learning and negatively impacted users' engagement with content. For tasks and controls that can be designed based on an image schematic input action, users performed most accurately with the metaphoric design. There are benefits and limitations to each approach to designing to support intuitive interaction. We conclude with preliminary design considerations, suggest ways to balance performance with high user satisfaction depending on contextual design goals and question a single definition of intuitive intuition within whole body interface design.**

RESEARCH HIGHLIGHTS

- We investigated three distinct strategies (metaphoric, isomorphic and 'everyday' or conventional) to designing intuitive interactional mappings.
- We compared the three mapping strategies based on measures of usability, intuitiveness and engagement.
- The metaphoric design enhanced users' performance completing tasks, but the lack of discoverability of the interaction model left them feeling incompetent and dissatisfied.
- The isomorphic design enabled users to focus on tasks rather than learning how to use the system.
- We provide preliminary guidelines on the benefits and limitations of each mapping strategy and when it would be ideal to use each.
- We provide the groundwork for future research that can do further comparisons on these mapping strategies.

*Keywords: natural user interfaces; tangible computing; metaphor; image schema; intuitive interaction; design guidelines*

*Received 7 May 2014; Revised 17 December 2014; Accepted 2 January 2015*

## 1. INTRODUCTION

Natural user interface (NUI) has become an umbrella term used to describe a variety of interfaces that use gestures and/or other body movement as input to control the system (O'Hara *et al.*, 2013). NUIs present three challenging design problems. First, there are very few standards or conventions that outline how certain controls for functions or tasks should be designed. This is largely due to the novelty of these interfaces.

Second, the supported interaction of NUIs is similar to actions people would do with physical environments, tools or devices in their everyday activities. It is difficult to determine *a priori* if users will try to interact as they might with other interactive devices or with 'everyday' mechanical or physical objects and/or environments. Third, the lack of physical or perceptual affordances in many whole body systems (in particular those that are based on camera vision) makes it difficult for users to know what gestures and actions are supported unless they are given very detailed instructions. In typical NUI systems, users can move in a variety of ways but a small subset of these actions are supported by the system. In a recent study by Hornecker (2012), children attempted 3D actions and actions based on their naïve understandings of various laws of physics on a tangible interface only to find out they were not supported. While Hornecker (2012) states that it is difficult to know what a user will expect from a system, a lack of affordances can actually foster learning and reflection—actions that are desirable in certain contexts.

These challenges not only make NUIs difficult to learn for users but leave very little design guidance for designers on how to make these interfaces usable and enjoyable. In an article that discusses the benefits and limitations of gestural interaction within human computer interaction (HCI), Norman (2010) states that interaction techniques such as gesture, touch and speech come with new problems and challenges that can potentially lead to mistakes and confusion in the user experience.

In this paper, we report the findings of a comparative study between three different interaction models for the same whole body system. Each model leverages a different approach to mapping body actions to system controls. We measure the usability and intuitiveness of each model as well as explore how each model affects the users' engagement with the content domain of the whole body system, which is social justice. We describe the three design strategies in detail, for other designers to use and compare with their own practice. We provide a methodology for measuring intuitive interaction that is based on previous research within HCI and Cognitive Science. From the results of our study, we provide guidelines that begin to address the challenges of designing for NUIs, which bring us closer to establishing set conventions on how to make these interfaces actually feel 'natural'.

## 2. SUPPORTING INTUITIVE INTERACTION

One approach to compensate for the lack of affordances NUIs present is to create control mappings that support *intuitive interaction*—interaction that is fast, unconscious and automatic. Bastick (1982) describes intuition as a cognitive process that uses information previously perceived by the senses. This subconscious use of previous knowledge gives users the ability to successfully use the interface almost instantly and with minimal conscious effort. Within the HCI and design literature, there are various interpretations and mechanism cited for what makes interaction intuitive (Antle *et al*., 2009b; Blackler *et al*., 2002; Hurtienne *et al*., 2008; O'Brien *et al*., 2008), which we further discuss below. We focus on three distinct strategies, each based on a different cognitive mechanism that designers can use to design interactional mappings that are intuitive. We present the three strategies below and show how each supports intuitive interaction.

### 2.1. Metaphoric mappings

*Metaphoric mappings* base input actions on image schemas—mental models formed from repeated patterns in everyday experiences—and system effects on related conceptual metaphors. A simple example is the metaphorical association of the image schema UP-DOWN with quantity. 'Up' is associated with 'more' and 'down' with 'less'. When we fill a cup or add objects to a pile, we notice the substance or object growing in height. The metaphor UP IS MORE is a cognitive structure based on these everyday experiences and used—unconsciously—to understand a variety of more abstract concepts. For example, we use this metaphor to make sense of system controls (e.g. raising the sound volume by moving the slider up). Figure 1 illustrates the concept of a metaphoric mapping.

Various researchers in HCI have explored the use metaphoric mappings in various tangible and whole body systems (Antle *et al*., 2009a, b; Bakker *et al*., 2011; Holland, 2010; Hurtienne *et al*., 2010; Svanaes, 2001). Some were interested in applying metaphoric mappings in abstract domains such as sound manipulation (Antle *et al*., 2011; Bakker *et al*., 2011; Holland, 2010). Others explore the use of metaphoric mappings as a design tool that can help the usability of

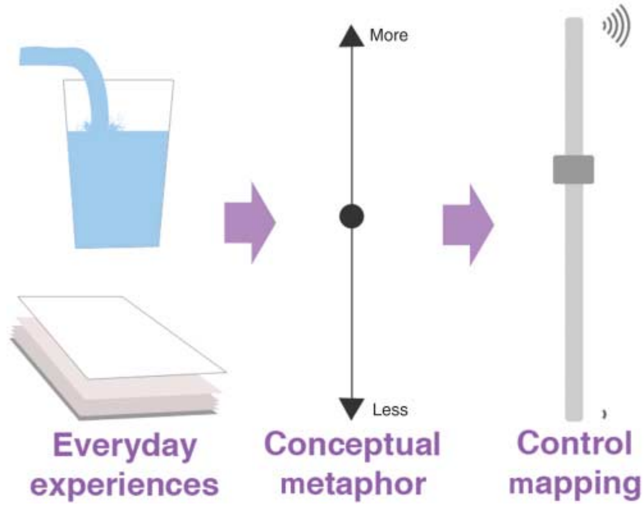


**Figure 1.** The underlying theory behind metaphoric mappings.

an interface (Antle *et al.*, 2009a, b; Hurtienne *et al.*, 2010; Svanaes, 2001).

Metaphoric mappings support intuitive interaction because the conceptual metaphors that frame the mapping are understood below the level of conscious awareness. Because of this, we call this cognitive process 'intuitive' and interaction based on it, 'intuitive interaction'.

### 2.2. Isomorphic mappings

*Isomorphic mappings* are one-to-one literal spatial relations between input actions and resulting system effects. The most common form of isomorphic mapping is physical–physical. An example is a racing game for a whole body system where the player's movement is mapped to a car's movements. Players turn their waist left to turn the car left. However, physical–physical mappings may not be possible in more complex systems. Another form of an isomorphic mapping is physical–abstract. For example, to control quantity of sound volume, one could map the height or amplitude of a sound wave (i.e. volume, measured in Decibels) to a horizontal line of filled in ticks. Each tick represents a set quantity or constant amount of sound volume. The system's sound volume is isomorphically mapped to the amount of filled or selected ticks in the horizontal line. To increase the sound volume, the user selects more ticks. For both examples, the input and system response have the same—isomorphic—structure. Figure 2 illustrates these two examples.

Smith (1987) describes physical–physical isomorphic mappings as literal and states that their key advantage is that they are very easy to learn. Both types of isomorphic mappings can be intuitive if the user understands the nature of the structure being controlled by the interaction. For example, the array of ticks may not be intuitive for a user who does not think of volume as a parametric value that could be increased at a constant rate.

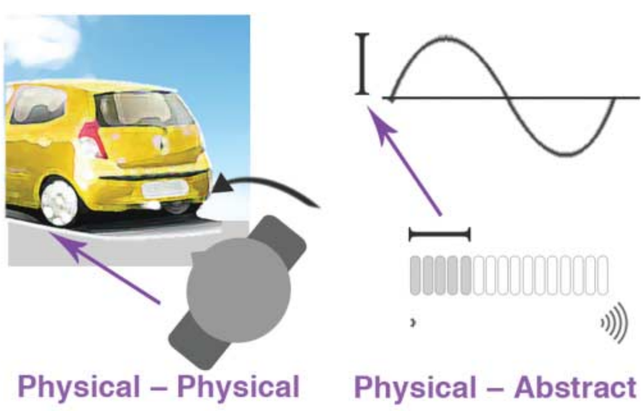


**Figure 2.** Examples of the two types of isomorphic mappings.

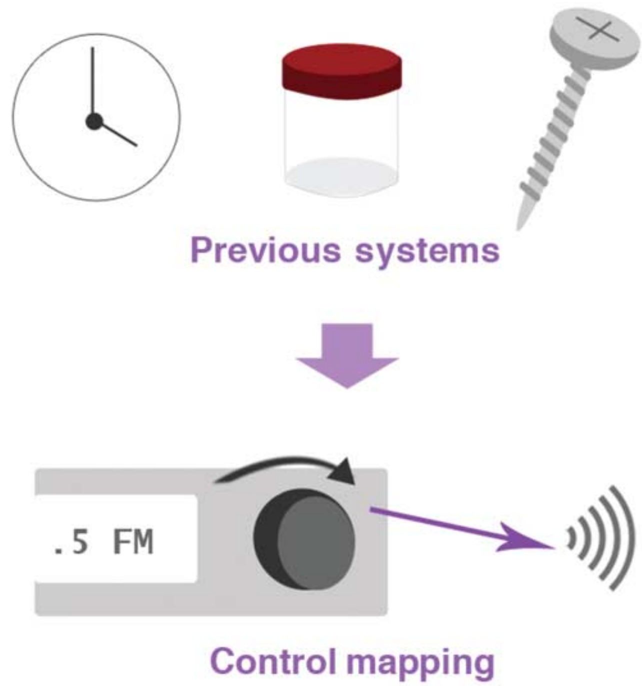


**Figure 3.** The theory behind conventional mappings.

### 2.3. Conventional mappings

We define *conventional mappings* as those adapted from previous practice and commonly found in product interfaces. When conventional mappings are found across a variety of interfaces, they become familiar to users (Blackler *et al.*, 2002). In order to differentiate conventional from metaphoric and isomorphic mappings, we limit them to those found in other systems but NOT grounded on image schema-based metaphors or one-to-one mappings. However, in most cases, their origins or structuring may be random. An example of such a mapping is the arrangement of letters on a QWERTY keyboard. Typically, conventional mappings have to be learned and take time to become established in design practice (Norman, 2010). An example of a conventional mapping for sound control is a physical dial that increases volume with a clockwise rotation. Associating clockwise movements with increased quantities comes from our experience with clocks, radio dials, screws and jars—clockwise rotations increase time, numeric values and tension. Figure 3 illustrates the theory behind conventional mappings.

Blackler *et al.* (2002) describe intuitive use as being based on our experiences with previous systems. In a comparative study between familiarity with features and functions used in technical interfaces and task completion, they found that people who had a higher familiarity with the interface features and functions had higher task completion rates and showed more instances of intuitive interaction (Blackler *et al.*, 2002). Conventional

**Figure 4.** The whole body system used in our study. Participant is using the isomorphic mapping design.

mappings can be intuitive when they are based on familiar experiences with other systems. However, the structures of these mappings may be arbitrary. Currently, very few conventional mappings exist for NUIs (Norman, 2010).

Given these three strategies, we had one over-arching research question and three hypotheses:

**Research question:** How does each mapping strategy compare in usability, intuitive interaction, awareness and impact?

**Hypothesis 1:** We hypothesized that the metaphoric mapping would have better usability than the other two mapping strategies due to previous work where systems using metaphoric mapping had high usability ratings (Hurtienne *et al*., 2008) and instances of successful interaction (Antle *et al*., 2011).

**Hypothesis 2:** We hypothesized that there would be a significant difference between the three strategies and how they support intuitive interaction. Although all strategies support intuitive interaction differently, the degree in which they do so is unknown.

**Hypothesis 3:** We hypothesized that the metaphoric mapping would have more significant effects on awareness and impact due to previous work that had high post-task awareness and impact ratings for a system using metaphoric mappings (Antle *et al*., 2011).

## 3. METHODOLOGY

We used an exploratory, comparative design with three conditions: metaphoric, isomorphic and conventional. We used a between subjects design to avoid learning and carry-over effects. Participants completed tasks on the same whole body system, called *Springboard*. Controls for the system were the same for each interaction model. Only the input actions varied between versions based on the three different interaction models (metaphoric, isomorphic, convention). We collected quantitative data including task completion times and Likert values for responses to survey questions. We also collected qualitative data including video, observational notes and responses to open interview questions.

### 3.1. Research prototype: springboard

For replication purposes, we used a system, called *Springboard*, which has been used in other research on intuitive interaction (Antle *et al*., 2011). Springboard is a whole body installation based on the abstract concept of balance in social justice. Social justice can be defined as the balanced or equitable distribution of the advantages and disadvantages within a society. Springboard supports users to interactively explore and reason about digital images related to three social justice issues: the distribution of food; the resources used for shelter and community control and safety. Each issue involves consideration of many factors. We simplify each issue to the consideration of two main factors which when balanced result in an equitable or socially fair solution. For details of the images used in Springboard and the general system architecture, see Antle *et al*. (2011).

We used the three strategies to design three new and distinct input-control mappings for Springboard. In general, this iteration of Springboard is designed to take as input, a user's position in space and based on this, displays different image pairs that represent different states of food management (Antle *et al*., 2011) (Fig. 4). This state could be balanced or imbalanced. For example, if a user is at the far edge of a permissible input space, then this triggers the display of images representing an

**Figure 5.** A top–down view of our metaphoric mapping with a user in an imbalanced state. The user is slightly left and sees images of higher environmental preservation and low food production. The pair of squares illustrates different states in food management. Pairs that have the same shade represent a balanced state.

imbalanced situation related to the social justice issue. Moving back into the center of a space enables the user to trigger imagery that depicts more balanced states related to that issue. The system can be customized with different interaction models by varying how the input space sensing is defined and mapped to image display controls.

The metaphoric mapping redesign is informed by a spatial representation of the twin-pan and point balance image schemas. Both schemas use balance metaphorically to describe the relationship between two different factors. The input space is the rectangular floor marker. The user's spatial position on the marker can be balanced (centered) or imbalanced (off-center, as illustrated in Fig. 5). When the user stands in the middle of the rectangle (i.e. the balance point), there is equal space or emphasis on the left and right side of him with respect to the rectangular space. That is, if the user moves to a spatially balanced position, the system displays images that depict *metaphorical* balance between the left and right images (i.e. twin-pan). Being more on the left or right side of the marker puts more emphasis to one side of the rectangle and creates an imbalanced state between the left and right images.

The isomorphic mapping redesign relates the user's position within two triangles to the content displayed by the images (Fig. 6). Each triangle represents the construct illustrated by the left or right image. The vertical height of the triangle section that the user is standing on is directly mapped to that triangle's respective construct amount. On the other triangle, the area adjacent to the user's position represents the other construct's amount. The area of the triangles is a direct representation of the conceptual relationship between the left and right image. Thus, the input and display space is isomorphic.

In both the metaphoric and isomorphic mappings, the center of the input space represents a balanced state. However, the remaining parts of each interaction model and the use of the interface are different according to a metaphoric or isomorphic model. During the experiment, we also asked participants to describe their understanding of the system and inferred if their mental model matched the system they were assigned to.

The conventional mapping redesign relates the user's position along a circular path laid out on the floor to the different states of food management (Fig. 7). Different areas on the path are mapped to different (im)balance states which increase or decrease linearly. Due to the few conventions available for whole body systems, we are leveraging a convention from another domain. This specific convention has transitioned successfully from analog (i.e. physical dial on radios) to gestural devices (i.e. navigational wheel on old MP3 players) and may transition well into a whole-body domain.

### 3.2. Participants

Thirty-two adults (13 males, 19 females) from the greater Vancouver area in Western Canada volunteered to participate in this study. Their age ranged from 18 to 55 years ($M = 26.9$, $SD = 8.3$). Seventy-two percent were university students (15 undergraduates, 8 graduates). Six percent were in their last year of high school. The remaining 22% had degrees and were working in industry. Twenty-four of participants used a computer and a smart phone daily. Others simply used a computer or a conventional cell phone daily. Twelve participants used tablets (i.e. iPad) daily or weekly. Ten participants were

**Figure 6.** A top–down view isomorphic mapping used in our study. The user is standing in the largest area of the top triangle and sees images with the highest environmental and lowest food production.

**Figure 7.** A top–down view of the conventional mapping used in our study. The user is at 6 o'clock and sees equal environment and food production.

randomly assigned the metaphoric condition, 11 the isomorphic condition and the remaining 11 the conventional condition.

### 3.3. Tasks

There were five task sets for the content set depicting the balance in the relationship between environmental preservation and food production. Each task set is associated with a reference code (e.g. T1 for task set 1). The task sets increase in difficulty from easy (i.e. 'Please explore how your movements affect the images on the screen') to hard ('Please show a sequence of moderately high food production, minimal environmental preservation, balanced food management and moderately high food production'). In total, participants completed a total of 10 tasks (1 in the first task set, 1 in the second set, 3 in the third set, 3 in the fourth set and 2 in the fifth).

T1, Exploration, can be thought of as a low-risk exploration period where the participant can familiarize herself with the system. Participants were given 5 min to explore the interface and observe how their movements affected the images on the screen. Questions regarding the interface were not answered because it was important to see how the participant understood the system with the given affordances and no instructions. Once a participant felt comfortable completing tasks using

**Table 1.** Variables associated with the Usability construct and the approach taken to analyze them.

| Variable | Collection method | Data type | Analysis method |
|---|---|---|---|
| Effectiveness | Task score | Ratio | ANOVA |
| Efficiency | Completion time | Ratio | ANOVA |
| User satisfaction | SUS | Ordinal | Kruskal–Wallis test |
| Self-perception of competence | PCS | Ordinal | Kruskal–Wallis test |

Springboard, she could begin the next task. Otherwise, she was told when 5 min was up. Participants did not have to explain their observations but show it in T2. Therefore, T1 did not have a success score.

In T2, Find Balance, participants were asked to make the left image and right image show equal states of food management. There was no time limit for this task. Participants had to tell the experimenter when they thought they had completed the task. After this task, participants were asked how they would teach a friend to use Springboard as well as what they thought the images represented. This task and the two interview questions measured their initial mental model and understanding of the system.

In T3, Show States, participants were asked to show specific states on the screen. An example task from this set is ‘Please show an above average amount of environmental preservation and below average food production’. They were told that more than one image could represent a state and that they did not need to look for a specific image. The experimenter also explained what was meant by environmental preservation and food production to avoid misinterpretations about the question. Participants were told to indicate when they had completed the task. Participants were asked to show three different states in total. The state and the order in which they had to display them were randomized.

In T4, Relative Change, participants were asked to go to a starting location in the input space. Starting from this location but being able to move, participants were asked to increase or decrease the amount of environmental preservation or food production. An example task from this set is ‘From your current position, please increase the level of food production’. Participants did this a total of three times. Although the specific location and order were randomized, each participant had to start in a position of perfect balance, a position where environment preservation dominated food production and a position where food production dominated environmental preservation.

In T5, Sequences, participants were asked to show a four-part sequence of states. Participants were instructed to indicate when they achieved a part of the sequence before moving to the next part. To ensure that participants were more focused on showing the sequence as opposed to memorizing it, they could ask the experimenter to repeat the next part of the sequence if they forgot it. Since this was the most difficult set of tasks, participants only needed to show two sequences. One sequence only included different levels of either environmental preservation or food production. An example of this type of sequence is ‘Please show us minimal environmental preservation, balanced food management, minimal environmental preservation, moderately high environmental preservation’. The other sequence included different levels of both. An example of this is ‘Please show moderately low environmental preservation, excess food production, balanced food management and moderately low environmental preservation’. This was done to see if they could think of the constructs independently and if sequences that only focused on one construct were easier to do. Each sequence included one repetition of a previous state as well as the perfect balance state.

The whole study took ∼45–60 min to complete. Participants were compensated for their time with \$10 in the form of cash or a gift card from the local coffee shop.

## 3.4. Measures

### *3.4.1. Usability*

We followed ISO 9241’s definition of usability (ISO 9241-11, 1998) and measured effectiveness, efficiency and user satisfaction (Table 1). We measured effectiveness by the amount of tasks a participant did correctly and converted this number into a percentage. Efficiency was defined as the mean time taken to complete each task and was measured in seconds. We measured user satisfaction with the system usability scale (SUS) (Brooke, 1996). The SUS is a 10-item Likert scale that measures a user’s feelings toward a system. Along with these three measures, we decided to also measure feelings of competence using the perceived competence scale (PCS) (Deci and Ryan, 1985). The PCS is a six-item Likert scale that measures the user’s feelings on how well they completed tasks using the interface. We added the fourth construct because previous studies have shown a positive correlation between users who feel confident in completing tasks and satisfaction with using a system (Deci and Ryan, 1985).

### *3.4.2. Intuitive interaction*

We measured intuitive interaction using four constructs: perceived intuitiveness, expectation, conscious attention and subconscious actions (Table 2). Perceived intuitiveness and expectation were included to account for Spool’s (2005) interpretation of intuitive use. He describes intuitive use as instances when the interface behaves as we expect it to (Spool,

**Table 2.** Variables associated with the Intuitive Interaction construct and the approach taken to analyze them.

| Variable | Collection method | Data type | Analysis method |
|---|---|---|---|
| Perceived intuitiveness | Likert scale | Ordinal | Kruskal–Wallis test |
| Expectation | Verbal answer | Nominal + qualitative | Descriptive statistics<br>Thematic analysis |
| Conscious attention | Verbal answer, Likert scale | Qualitative + ordinal | Triangulate with effectiveness, Kruskal–Wallis test |
| Subconscious action | Video recording | Qualitative | General observation |

**Table 3.** Variables associated with the Engagement construct and the approach taken to analyze them.

| Variable | Collection method | Data type | Analysis method |
|---|---|---|---|
| Awareness | Likert scale | Ordinal | Kruskal–Wallis test |
| Impact | Likert scale | Ordinal | Kruskal–Wallis test |

2005). We measure perceived intuitiveness with a self-reported 7-degree rating on how intuitive they found the system after completing all the tasks of the study. We assessed expectation through an open-ended interview question at the end of the study. We asked participants if the interface behaved as expected and to explain the reasons behind their answer.

Conscious attention and subconscious action refer to Bastick's (1982) and Lakoff and Johnson's (2003) views on intuition within Cognitive Science. Both Bastick (1982) and Lakoff and Johnson (2003) suggest that we use cognitive structures formed from repeated sensory-motor experiences to make sense of—or intuit—new experiences or situations. This process occurs below the level of conscious awareness. When an interface is designed to support the subconscious nature of this cognitive process, users can use their conscious attention to understand and complete a task as opposed to learning how to use the interface. In our study, we measure participants' conscious attention in two ways. First, we determined how well they consciously understand how the system worked. After T1 (Exploration) and T2 (Find Balance) were complete, we asked two questions. The first question ('If you were going to teach someone how to use this system, what would you say?') provided information about the participants' understanding of the system and its controls. The second question ('What do the images represent?') provided information about the participants' understandings of the content within the system (i.e. the relationship between food production and environmental preservation). The participants' explanations revealed their conscious and explicit understanding of the controls and content. We then combined this with their performance accuracy on T2 (Find Balance). A high task score was categorized as 80% or higher. A combined high score indicated that they could verbally explain and physically demonstrate how the system worked, which implies that their conscious attention could then be focused on the task. Second, after each T2 through T5 we asked participants to rate how they felt their attention was focused (where 1 = on using their bodily movement to use the system, 7 = on solving the problem involved in the task).

We measured participants' subconscious actions by observing which gestures and body movements the participants did when they were verbally describing how the system worked. Previous work has shown that participants who used interface designs using metaphoric mappings subconsciously moved parts of or all of their bodies in ways that mimicked the image schemata used as the basis for the controls of the interface even when they had difficulty in verbally describing how the controls worked (Antle *et al*., 2011). For example, when asked to describe how a previous version of the system used in this study worked, several participants said they were unsure while moving their hands up and down like a scale or twin-pan balance. We looked for this type of subconscious expressions by participants of the balance image schema.

*3.4.3. Engagement*

In our study, we measured engagement through measures of awareness and impact as per Antle *et al*. (2011) (Table 3). Awareness was defined as the user's knowledge toward issues of social justice. In this context, social justice referred to the ways we develop and maintain a community in terms of housing, food and security. Impact referred to the user's willingness to participate in improving issues within social justice. Both awareness and impact were self-reported 7-degree ratings taken before and after using the whole body system. We were interested in seeing how the different designs affected the user's understanding and feelings toward the issues explored by the images of the system. By exploring this, we could better understand the different mapping strategies' effects on topic awareness and impact—an important element for systems geared toward education, serious games and art installations.

## 3.5. Data collection

Prior to the study, participants completed a short survey that took their demographics, technology use, and initial awareness and impact measures. During each task, the system recorded task completion times and their final (im)balance state. We recorded the whole experiment on video for further analysis after the study. After each task set, participants were asked to rate the distribution of their attention (our second measure of conscious attention). In between the first and second tasks sets, we asked participants two interview questions about their explicit understanding of the system (our first measure of conscious attention). We asked, 'If you were going to teach someone how to use this interface, what would you say?' We analyzed responses to this question to identify and compare participants' explanations and descriptions of how they thought the controls worked, with how they actually worked. We also asked, 'What do the images represent?' We analyzed their responses to see if they understood how the two images were related (i.e. that more environmental preservation was linked to less food production, and vice versa). These two interview questions were scored. Each question was scored out of two: zero for no understanding, one for partial understanding and two for good understanding. We summed these two scores in order to represent each participant's overall understanding of the system controls and abstract concepts. A score of 4/4 represents strong understanding, three represents good understanding, two represents partial understanding, one represents poor understanding and zero represents no understanding of the system controls or content. This method of quantifying understanding from verbal data was inspired by previous research from Antle *et al*. (2009b) that used similar questions and scoring system to measure a user's understanding of Sound Maker. After the fourth task set, participants participated in a final interview where they explained how well the system met their expectations. For example, we asked them, 'Did the interface behave as you expected it to? If not, what did you expect it do?' We analyzed responses to identify their expectations and if they system worked as expected or not, and if not, in what ways their expectations were not met. After the study, participants filled out a questionnaire that measured perceived intuitiveness, user satisfaction (SUS), perceived competence (PCS) and post-task awareness and impact.

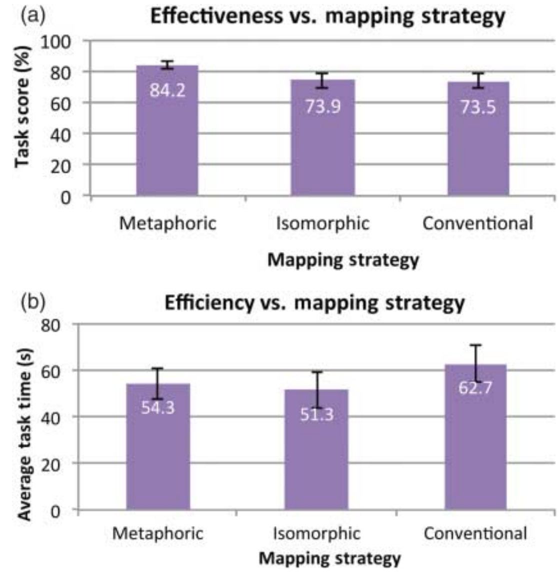


**Figure 8.** Mean task completion scores (**a**) and times (**b**) across the different mapping strategies. Whiskers represent standard error.

## 4. RESULTS

### 4.1. Usability

We ran inferential statistics on the four constructs of usability to see if mapping strategy had an effect on usability and if metaphoric mappings in particular had better usability than the other two mapping strategies (Hypothesis 1). Overall, we did not find significant difference on usability between the mapping strategies. However, we found significant differences on individual constructs of usability and interesting relationships between the usability constructs themselves. In regards to our hypothesis, metaphoric mappings did not have a significant effect on usability but generally had better task accuracy scores for certain tasks and smaller task completion times. To our surprise, metaphoric mappings also had generally lower satisfaction ratings in comparison with the other mapping strategies. We provide the detailed findings to each construct below.

#### *4.1.1. Effectiveness and efficiency*

We calculated the mean task scores (Fig. 8a) and completion times (Fig. 8b) across the three conditions. A one-way between subjects ANOVA did not show any significant differences between effectiveness or efficiency and the mapping strategy. However, further analysis and observational notes at an individual task level (vs. aggregated), revealed some key differences relevant to the research questions. The most striking was that *all* users in the metaphoric group got T2 (Find Balance) correct. Statistically, this was reflected in the finding that metaphoric mappings had significantly higher T2 scores than isomorphic mappings ($F(2, 29) = 9.021$, $P = 0.011$, $\eta^2 = 0.451$). Another interesting finding was that participants in the conventional group, on average, took 10 more seconds to complete all of the tasks. Our interpretation of observational data was that users in the conventional group took longer due to the time spent initially learning how to use the interface for each task set. Lastly, all users quickly completed T4 (Relative Change) and our interpretation of observational notes and open interview question data was that this efficiency was made possible by drawing the participant's attention to their location in the input space, which was done for all participants at the starting point for this task.

#### *4.1.2. User satisfaction and perceived competence*

We calculated the median SUS (Fig. 9a) and PCS (Fig. 9b) scores across the three mapping conditions. Kruskal–Wallis tests showed no significant differences between user satisfaction (SUS) or perceived competence (PCS) and mapping strategies. However, further analysis revealed an interesting relationship between the measurements themselves. First, correlation analysis revealed that SUS and PC scores were strongly correlated, that is, users who felt competent using the system also felt satisfied (Spearman's $r = 0.802$, $P < 0.0001$). Second,

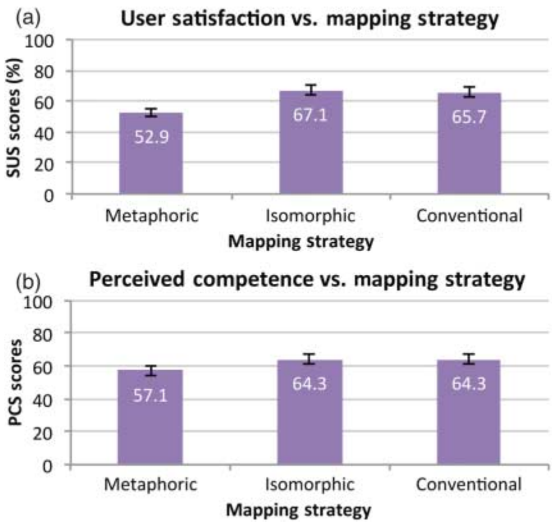


**Figure 9.** SUS (**a**) and PCS (**b**) scores across the different mapping strategies. Whiskers represent percentage error.

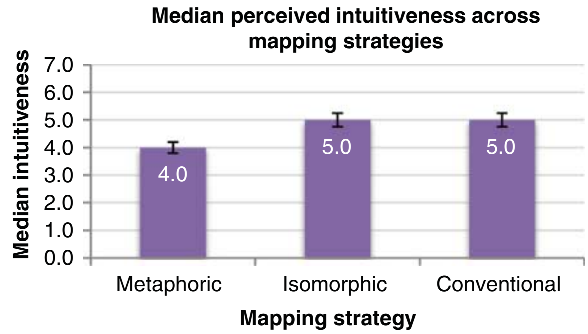


**Figure 10.** Median perceived intuitiveness ratings across the different mapping strategies. Whiskers represent percentage error.

we found that, on average, all the SUS scores were low, not exceeding the 70% mark. Third, we found that participants in the metaphoric group (Med = 37/70), on average, gave satisfaction scores that were 10% lower than the other groups (isomorphic Med = 47/70; conventional Med = 46/70). This contrasts with their efficiency and effectiveness scores. We also saw this contradiction in the qualitative interview data. For example, when asked if the system met their expectations, one participant said, 'No it did not, I expected left is more nature, right is more waste.' (U7, metaphoric condition). This quote is interesting because the user's task scores indicated that they used the system correctly. The quote also demonstrates the proper metaphoric model. U16 (metaphoric), who also had a high task score, was discussing mismatched expectations while sidestepping, left, right and back to center position. Her body motion reflected the balance conceptual metaphor even though her expectations were not met. Both U7 and U16 demonstrated high task scores and some degree of subconscious understanding. However, they stated that the system did not meet their expectations. We suggest that perhaps users' subconscious understanding of the system did not meet their conscious expectations and resulted in an unsatisfying experience.

## 4.2. Intuitive interaction

We ran descriptive and inferential statistics to see if mapping strategy had a significant effect on intuitive interaction (Hypothesis 2). Overall, mapping strategy did not have a significant effect on the instances of intuitive interaction. We did find a significant difference between mapping strategies and a construct of intuitive interaction: conscious attention. We also describe instances of intuitive interaction based on the relationship between a participant's conscious attention and subconscious actions. We provide the detailed findings for each construct below.

### 4.2.1. *Perceived intuitiveness*

We calculated the median intuitiveness ratings across mapping conditions (Fig. 10). A Kruskal–Wallis test showed no significant difference between perceived intuitiveness and mapping strategies. While the median intuitiveness ratings were not significantly different, our observational notes and analysis of users' responses to open interview questions suggested that participants acted and described their 'intuitive interactions' in different ways. For example, while it took participants in the conventional group slightly longer to complete tasks, once they consciously determined or learned the mapping, they easily applied it to related subtasks. This may be in part because the circular design of the input space, combined with previous usage of dials in a variety of devices, provides strong cues about how to move to a new spatial location on the ring. We also found that over half of the participants in the metaphoric group who demonstrated poor ability to verbally explain how the system worked, had high accuracy on task scores.

### 4.2.2. *Expectation*

We calculated the distribution between participants who felt that the interface design met their expectations and those who did not (see Fig. 11) across groups. We found an even distribution between participants who felt the interface met their expectations and participants who did not feel that the interface met their expectations with no clear pattern between groups. However, we find it interesting that half of participants did not have their expectations met. We identified and grouped reasons participants reported about why the system did not meet their expectations. Our analysis revealed three key themes to their unmet expectations:

(i) Unhappiness with controls of the system ($n = 4$).
(ii) A desire for more control ($n = 6$).
(iii) Difference between perceived and actual input-control mapping ($n = 6$).

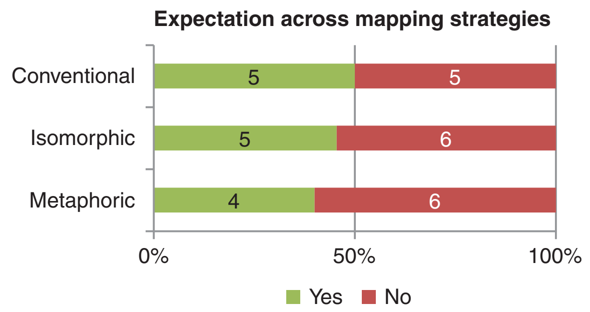


**Figure 11.** Distribution of participants who (did not) feel that the interface design met their expectations.

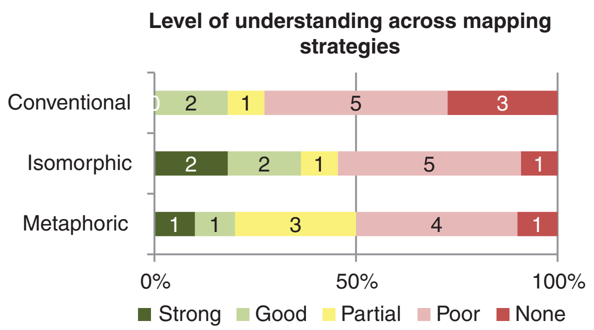


**Figure 12.** Distribution of participants who understood the system completely, partially and not at all.

### 4.2.3. *Conscious attention*

We calculated the distribution across mapping strategies of our first measure of participants' conscious attention. That is, how well participants could verbally explain and physically demonstrate (through accurate performance) that they understood how the system worked. We had five categories based on combining the verbal score (/4) and the accuracy score (%). Participants were classified as: strong, good, partial, poor and none, in terms of their understanding of the system after T2. We assume participants that understand the system well can devote conscious attention to the task rather than learning how to use the system. We compare 'understanding' as an indicator of conscious attention across mapping strategies (Fig. 12). Values suggest a fairly even distribution across mapping strategies.

For our second measure of conscious attention (self-ratings), we calculated the median score (1 = attention on their movement to use the system, 7 = attention on solving the task) across all mapping conditions (Fig. 13). We ran a Kruskal–Wallis test to see if there were any significant differences between where a participant placed their attention and the mapping condition. We found a significant difference between the groups ($\chi^2(2, 29) = 9.327$, $P = 0.009$). Post-hoc comparisons using a pairwise comparison Mann–Whitney test with Bonferroni correction show a significant difference between the attentional ratings of metaphoric (Med = 3.75) and isomorphic (Med = 5.5) mappings ($P = 0.015$). No significant differences were found between conventional mappings (Med = 4.0) and the other two mapping strategies.

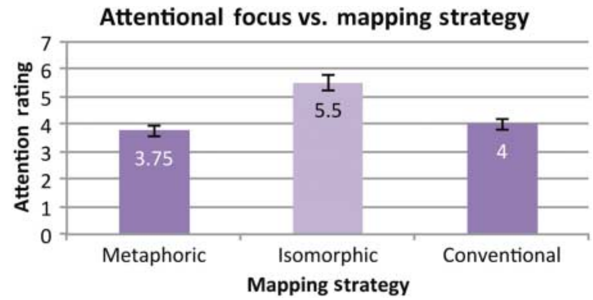


**Figure 13.** Attentional focus ratings across the different mapping strategies.

Detailed analysis comparing measures revealed that six participants had partial, poor or no understanding, high tasks scores and attentional ratings focused on completing the task (i.e. 5.0 or greater) rather than on using movement to interact with the system. We propose that these individuals may have used intuitive interaction because their task scores were high and they reported focusing on the task but they were not able to accurately explain how the system worked, suggesting a subconscious understanding of the system. A subconscious understanding of the system (rather than conscious or explicit), enabled them to focus their conscious attention on completing the tasks, not on learning to use or using the interface.

## 4.3. Engagement

We ran inferential statistics on the two constructs of engagement to see if there was a significant difference on engagement between the mapping strategies and if metaphoric mappings in particular had higher impact and awareness scores (Hypothesis 3). Overall, we did not find a significant difference between the mapping strategies and engagement. However, we did find significant differences between the pre- and post-impact scores for the conventional mapping. We describe the exact findings below.

We calculated the medians of the pre- and post-ratings of awareness and impact for each mapping condition. Wilcoxon tests showed no significant differences for the change between the pre- and post-awareness or impact ratings for the metaphoric or isomorphic groups. However, there was a significant decrease between pre- and post-impact ratings for the conventional condition ($W = 47.5$, $P = 0.044$, $\eta^2 = 4.75$). Kruskal–Wallis tests showed no significant differences between post-awareness or post-impact ratings and mapping strategies.

During the post-task interview, one participant stated that the embodied nature of the interaction made him feel more engaged with the material. He continued to state that because his movement matched the model of the concepts being explored, it was more effective in portraying a message than having someone say it to him.

The connection between body movement and the themes made a stronger impression […] If you consider politics or something, you get left/right associations with things. Having an interface where your body is engaged to pre-set associations to movement is more impressionable and memorable. (U9, conventional).

## 5. DISCUSSION

### 5.1. Usability

Metaphoric mappings had significantly higher T2 scores than Isomorphic mappings. However, all mapping strategies had similar effectiveness, satisfaction and perceived competence results. This contradicts our first hypothesis. These findings do not reflect the results from previous studies (Antle *et al*., 2011; Hurtienne *et al*., 2008). However, it is important to note that previous studies did not compare the same mapping strategies as this study.

There was a significant positive correlation between user satisfaction and perceived competence. Understanding the relationship between satisfaction and competence can provide insight on designing systems that are easy and satisfying to use. While metaphoric mappings resulted in higher effectiveness and good efficiency, users were feeling unsatisfied with their experience using the system. Furthermore, while participants assigned to the conventional mapping condition demonstrated lower efficiency, they were still fairly satisfied with their performance. We believe this relates back to the nature of metaphoric and conventional mappings. Metaphoric mappings are perceived by the senses and represent previous knowledge subconsciously used. Conventional mappings on the other hand are acquired through reflection and learning and represent previous knowledge that was consciously used. With metaphoric mappings, many who had high task scores still lacked an explicit understanding of how the system worked. However, based on task time and observational notes, we inferred that those who used the conventional mapping design seemed to take longer in order to learn how the system worked as the study progressed, even if their task scores were poor throughout.

These findings suggest two key concepts. The first is the importance of discoverability of the interaction model. This is consistent with previous work that suggests that discoverable mappings are needed to support intuitive interaction (Antle *et al*., 2009b; O'Brien *et al*., 2008). Furthermore, discoverable mappings help users understand what they can and cannot do within a system—which is especially important if the system lacks affordances to give this type of information (Hornecker, 2012). To make mappings more discoverable, designers should integrate tight mappings between input and control (Antle *et al*., 2009b; Hornecker, 2012) as well as provide salient feedback that hints at the possible system effects of input actions (O'Brien *et al*., 2008).

The second concept these findings suggest is the relationship between understanding the interaction model of a system explicitly and having a satisfying user experience. Users who can understand the interaction model will find the design behind the system clear, the functions better integrated and feel more competent about being able to use the system successfully. All these traits refer back to high user satisfaction (Brooke, 1996) and perceived competence (Deci and Ryan, 1985). To ensure a satisfying user experience, designers should create mappings that are literal (Smith, 1987) or leverage the user's knowledge of previous systems that have similar functions (Norman and Nielsen, 2010).

### 5.2. Intuitive interaction

The three mapping strategies were similar in how intuitive they felt and in their ability to foster an explicit understanding of the system (Fig. 12). However, it is interesting to note a slightly higher number of participants who felt as if the interface did not meet their expectations in the metaphoric mapping condition (Fig. 11). These findings suggest a difference between knowing how to use the system and knowing that you have used it successfully. Participants who were unaware of their successful performance subconsciously knew how to use the system but were consciously unaware of it. This correlates to previous work that had similar findings with a system using metaphoric mappings (Antle *et al*., 2011). Furthermore, in a study by Hornecker (2012), participants tried many actions that did not map out to successful system effects. These instances of unsuccessful interaction illustrate mismatched expectations and no knowledge of how the system works. We suggest that having knowledge of both the system and instances of successful interaction are important when meeting user expectations. Designers need to be make the interaction model easily understood by the user and provide clearer feedback on actions that are supported by the system.

Furthermore, participants who were assigned the isomorphic mapping condition placed more of their attention on performing the task and less on their body movement in comparison with participants who were assigned the metaphoric or conventional mapping condition (Fig. 10). However, since other constructs within our intuitive interaction measure showed more equal distributions across mapping conditions, we cannot confidently attribute these attentional ratings to intuitive interaction. Instead, we provide possible reasons behind the lower ratings in the other two conditions. In the metaphoric and conventional mapping conditions, participants were unclear on the interaction model and spent more time trying to relate their movement to the system effects. They spent a lot of time trying to learn the input-control mappings and to reflect on their observations from the previous tasks. These results suggest the importance of straightforward mappings in maintaining the user's attention on completing the task. While there are times when body focus is important (i.e. dance, surgery), the execution of movement should still fall in a semi-conscious state (i.e. automatic). Hornecker (2012) also discusses instances when intuitive interaction may not be desired and that designers may want users to reflect and learn the controls of the system. Knowing when systems should rely on automatic movement or when they should elicit reflection is an area that needs further research.

### 5.3. Limitations

Due to low sample size and a between-subjects design, we state our claims cautiously. Furthermore, to balance the three mapping conditions, interface designs were very similar and may not leverage the best aspect of a certain mapping strategy. Had we focused on making the perfect conventional, isomorphic or metaphoric example, we would have increased the differences between the interface designs and may have had more significant results. However, this would introduce confounds that could account for significant differences found in data analysis. In addition, our intuitive interaction construct relied on self-reports, which may not provide accurate reflections of subconscious use. However, self-report is a measure that is often used to measure this construct. Lastly, our research prototype is of a very specific context and application. Similar studies that compare the different mapping strategies in different contexts are needed.

## 6. DESIGN IMPLICATIONS

Based on our findings, we offer the following set of design considerations and preliminary guidelines (Table 4). These are preliminary and will need further exploration in future studies before they become established guidelines.

Here are two examples to illustrate how the above table could be used. A group of designers want to create an interactive color-mixing installation in a science museum. They understand the system will be in a walk-up-and-play environment and that the system may have a 5-min usage cycle. In this scenario, the designers want the children to focus on the task of color mixing rather than learning how to use the interactive system. They also want the children to reflect on the cause and effect of mixing different colors. Based on Table 4, they should consider either using an isomorphic or conventional mapping. Another group of designers would like to create a game to teach attention deficit hyperactivity disorder (ADHD) children simple meditation techniques. Based on this user group, the game must be mentally effortless as ADHD children may react negatively to mentally stressful situations. In this scenario, the designers should consider using a metaphoric mapping between user input and system controls.

**Table 4.** Design considerations based on the desired goal of the system.

| Desired goal | Design recommendation |
|---|---|
| Discoverable input actions and controls | Create literal mappings between input action and control while providing salient feedback |
| High usability | Pair metaphoric mappings with salient feedback to make instances of successful interaction explicit |
| High effectiveness and efficiency | Use metaphoric mappings |
| Automatic input, unconscious effort | Use metaphoric mappings |
| User attention focused on task rather than interaction with system | Use isomorphic mappings |
| System that matches user expectations | Use isomorphic mappings |
| Explicit understanding of system functions and high satisfaction | Use isomorphic or conventional mappings |
| User satisfaction | Use isomorphic or conventional mappings |
| Learning and reflection | Use conventional mappings |

## 7. CONCLUSIONS

We described an exploratory mixed-methods approach to compare three mapping strategies in terms of usability, intuitive interaction and engagement. While mapping strategy had a significant effect on usability and impact for certain tasks, overall there were no statistical differences between the different mapping strategies and usability, intuitive interaction and engagement.

Despite this, there were other observations that provided insights on the relationship between effectiveness and user satisfaction, as well as the relationship between body movement and content engagement. From the quantitative analysis, there was a significant difference between the mapping strategies and where participants placed their attention during the whole experiment. In particular, participants in the isomorphic condition gave attentional ratings that indicated more attention on completing the task then on using the system. The findings from this study replicate those of previous work (Antle *et al.*, 2009b, 2011; Hornecker, 2012; Norman and Nielsen, 2010; O'Brien *et al.*, 2008; Smith, 1987) and show areas that need further research.

By triangulating the results from the different research constructs, we present potential benefits, setbacks and uses for each mapping condition. Metaphoric mappings have the greatest potential to support automated learning of movement and subconscious learning of controls. Isomorphic mappings seem most suitable for walk up and play applications, which desire immediate use and minimal instruction. Conventional mappings demonstrated the greatest potential for learning applications, which have no literal mappings but can use analogies to create an explicit understanding of the action-control or action-meaning mapping.

We provide soft guidelines on the benefits and limitations of each mapping strategy and when it would be ideal to use each. Most importantly, we provide the groundwork for future research that can do further comparisons on these mapping strategies, explore the importance of intuitive interaction within HCI, and extend our methodology to other NUI contexts.

## ACKNOWLEDGEMENTS

We would like to thank all the participants in this study. We thank Gregory Corness and Martin Rittler for their technical help implementing Springboard.

## FUNDING

This work was made possible through funding by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Graphics, Animation and New Media Network Centre of Excellence, Canada (GRAND NCE).

## REFERENCES

Antle, A.N., Corness, G. and Droumeva, M. (2009a) What the body knows: exploring the benefits of embodied metaphors in hybrid physical environments. Interact. Comput., 21, 66–75.

Antle, A.N., Corness, G. and Droumeva, M. (2009b) Human-computer Intuition? Exploring the cognitive basis for intuition in embodied interaction. Int. J. Arts Technol., 2, 235–254.

Antle, A.N., Corness, G. and Bevans, A. (2011) Springboard: Designing Image Schema Based Embodied Interaction for an Abstract Domain. Human Computer Interaction Series. Whole Body Interaction, pp. 7–18. Springer.

Bakker, S., van den Hoven, E. and Antle, A.N. (2011) MoSo Tangibles: Embodied Metaphors in a Tangible System for Learning: A User Evaluation. TEI 2011, Funchal, Portugal. ACM Press.

Bastick, T. (1982) Intuition: How We Think and Act. Chichester, Toronto.

Blackler, A., Popovic, V. and Mahar, D. (2002) Intuitive Use of Products. Common Ground Design Research Society Int. Conf., London, UK. Staffordshire University Press.

Brooke, J. (1996) SUS: A Quick and Dirty Usability Scale. In Jordan, P.W., Thomas, B., Weerdmeester, B.A. and McClelland, I.L. (eds), Usability Evaluation in Industry. Taylor & Francis, London, pp. 189–194.

Deci, E.L. and Ryan, R.M. (1985) Intrinsic Motivation and Self-Determination in Human Behavior. Plenum, New York.

Holland, S. (2010) Asymmetrical Multi-User Co-operative Whole Body Interaction in Abstract Domains. CHI'10—Whole Body Interaction Workshop, Atlanta, Georgia, USA. ACM Press.

Hornecker, E. (2012) Beyond Affordance: Tangible's Hybrid Nature. TEI 2012, Kingston, Ontario, Canada, pp. 175–181. ACM Press.

Hurtienne, J., Weber, K. and Blessing, L. (2008) Prior Experience and Intuitive Use: Image Schemas in User-Centered Design. In Langdon, P., Clarkson, J.P. and Robinson, P. (eds), Designing Inclusive Futures. Springer, Berlin, pp. 107–116.

Hurtienne, J., Stößel, C., Sturm, C., Maus, A., Rötting, M., Langdon, P. and Clarkson, J. (2010) Physical gestures for abstract concepts. Inclusive design with primary metaphors. Interact. Comput., 22, 475–484.

International Organization for Standardization. (1998) ISO 9241-Part 11: Guidance on usability. ISO.

Lakoff, G. and Johnson, M. (2003) Metaphors We Live By. University of Chicago Press, Chicago.

Norman, D. (2010) Natural user interfaces are not natural. Interactions, 17, 6–10.

Norman, D.A. and Nielsen, J. (2010) Gestural interfaces: a step backwards in usability. Interactions, 17, 46–49.

O'Brien, M.A., Rogers, W.A. and Fisk, A.D. (2008) Developing a Framework for Intuitive Human-Computer Interaction. 52nd Annual Meeting of the Human Factors and Ergonomics Society, New York City, NY, USA.

O'Hara, K., Harper, R., Mentis, H., Sellen, A. and Taylor, A. (2013) On the naturalness of touchless: putting the "interaction" back into NUI. ACM Trans. Comput.-Hum. Interact. Special issue on embodied interaction, 20, Article 5.

Smith, R. (1987) Experiences with the Alternate Reality Kit: An Example of the Tension Between Literalism and Magic. SIGCHI/GI, Toronto, Ontario, Canada, pp. 61–67. ACM Press.

Spool, J.M. (2005) What makes a design seem 'intuitive'? User Interface Eng. http://www.uie.com/articles/design_intuitive/ (accessed August 19, 2011).

Svanaes, D. (2001) Context-aware technology: a phenomenological perspective. Hum. Comput. Interact., 16, 379–400.