Balancing Justice: Comparing Whole Body and Controller-based Interaction for an Abstract Domain

Alissa N. Antle*, Greg Corness, Allen Bevans

School of Interactive Arts + Technology
Simon Fraser University Surrey
250 -13450 102 Avenue Surrey, B.C.
V3T 0A3 CANADA
E-mail: {aantle; gcorness; alb19}@sfu.ca
*Corresponding author

Abstract: In this paper we present a quantitative, comparative study of a multimedia environment about social justice that users can control using whole body interaction or a simple control device. We explore the efficacy of using embodied metaphor-based whole body interaction compared to controller-based interaction for an abstract domain (social justice). We describe how conceptual metaphor theory can be applied to the design of a whole body interaction model, focusing on the twin-pan balance image schema and its metaphorical elaboration that structures the concept of balance in social justice. We describe the Springboard system, our methodology and results from a study with 76 participants. Our results indicate that participants were able to interact with our system using both input approaches. However, participants in the whole body group were more deeply impacted by their experiences related to social justice than those in the control device group.

Keywords: Embodied interaction, whole body interaction, embodied schema, image schema, metaphor, conceptual metaphor theory, social justice, interactive environments, movement interaction.

Reference for publisher use only

Biographical notes: Dr. Antle is an Associate Professor in the School of Interactive Arts + Technology at Simon Fraser University, Canada. Her research focuses on embodied human–computer interaction and child-computer interaction and proceeds through the design and evaluation of tangible and multi-touch surfaces, and interactive environments. Dr. Antle holds Bachelor degrees in Systems Design Engineering and Liberal Arts from the University of Waterloo, Canada and a Ph.D. from the University of British Columbia, Canada. Before returning to academe, Dr. Antle spent eight years in the new media industry working as a senior designer, executive producer and consultant.

Greg Corness is a researcher and artist working in media environments and interactive performance to explore embodied and social interaction with media agents. He publishes and lectures on electronic music, human-computer interaction and performance theory. His work on interactive systems includes generative sound systems, computer vision and tangible interfaces for galleries installations, interactive museum exhibits and live performance presented in Canada and the US. Greg holds a Masters of Music in

Computer Composition from the University of Victoria, Canada, and is currently completing his Ph.D. at the School of Interactive Arts + Technology, Simon Fraser University, Surrey Canada.

Allen Bevans is a Ph.D. student at the School of Interactive Arts + Technology at Simon Fraser University, Canada. He completed a Bachelor of Fine Art in New Media at the University of Lethbridge, Canada and a Master's of Science from the School of Interactive Arts + Technology. Allen's research focuses on design education, supporting creativity with technology, and interaction design for children. He is currently working on projects intended to help designers build better interactive systems for children.

1 Introduction

Whole body computing is increasingly becoming commercialized through devices such as the Nintendo Wii and Xbox Kinect. Little research exists on how to design effective interactional mappings between input actions and system responses for these whole body systems (Antle et al., 2009a; Svanaes, 2001). Often mappings are designed based on traditional controller design conventions or the designer's "intuition" (see Antle et al., 2009b) for a discussion of intuition in interaction design). A second issue is that research investigating what whole body interaction has to offer abstract application domains is largely overlooked territory (Antle et al., 2011; Holland, 2010). Image schema and conceptual metaphor theory offers an explicit motor-cognitive mechanism that structures thought in abstract domains based on physical action (Johnson, 1987). The theory explicates how image schemas are formed from physical experiences, and how conceptual metaphors based on these "embodied" image schema are derived. For brevity, we refer to this theory as *embodied metaphor theory*. In previous work, we have proposed, described and validated that embodied metaphor theory can be used to inform whole body interaction design for perceptual (Antle et al., 2009a; Antle et al., 2009c) and abstract domains (Antle et al., 2011). In these papers, we have described how the structural relationship between image schemas and metaphorical concepts can be mirrored in the computational structures that map input actions to output responses. We have also presented evidence from several different studies that this approach positively affects users' enactments, interactions and interpretations.

In this paper, we build on our previous research by presenting the results from a quantitative study in which we investigate how embodied metaphor-based whole body interaction compares to controller-based interaction for the same interactive environment. The interactive environment, called Springboard, enables users to explore images and sounds depicting various states of balance and imbalance in issues in social justice. For example, users can explore multimedia content about *balancing* sustainable and ethical agricultural practices with the quality and quantity of food produced to feed a growing population. The meaning of *balance* here is an abstract concept. Johnston suggests that our understanding of the meaning of balance in the abstract domain of social justice is understood through unconscious metaphorical elaboration of image schemas that pertain to balance in our bodies (Johnson, 1987). We build on this notion and explore how the metaphorical relation between body-balance and meaning-balance can be leveraged in whole body interaction design.

We begin this paper by summarizing other studies that have dealt with embodied metaphor theory in whole body and tangible interaction design research. Then we describe the theoretical foundations of embodied metaphor theory and explain how it can be applied to whole body interaction design. Next we describe the system implementation of Springboard. We present the methodology and results from a quantitative log and survey-based comparative study with 76 participants that enabled us to explore the following high level research question: How does embodied metaphor-based whole body interaction compare to using a simple controller for an abstract application domain? For brevity we abbreviated embodied metaphor-based whole body interaction as EM-WBI. We conceptualize interaction using a simple control device (e.g. a slider that controls the display of images) as a baseline, and compare EM-WBI to this baseline for a variety of constructs. Specifically, we address the following research questions:

Usability

1. Does incorporating an embodied metaphor-based whole body interaction (EM-WBI) model make the system more *efficient* and more *effective* to use?
2. Does an EM-WBI approach affect the *extent* of users' *exploration* of the multimedia content (i.e. do they explore the whole set of images and sounds or stop short)?
3. Does an EM-WBI approach affect users' *satisfaction* with their *performance* or ability to interact with the system?

Motor-Cognitive Processes

4. Are users consciously *aware* of the image schema instantiated in the system?
5. Does an EM-WBI approach affect users' ability to *focus* their attention on the multimedia content?

User Experience

6. Does an EM-WBI approach make the system more *enjoyable* or *interesting* for users?
7. Does an EM-WBI approach affect users' sense of *competency* using the system?
8. Does an EM-WBI approach enhance the *impact* of the experience?

We present and discuss the results from our quantitative comparative user study designed to explore these questions and constructs. In a separate paper, we report on a qualitative observational and interview study of users' experiences interacting with only the whole body implementation of Springboard. In that paper, we focus on identifying themes related to how users enact, interact, and interpret EM-WBI and explore differences between Springboard implementations using body-centric and spatial image schemas. In this paper, we discuss the similarities and differences in users' experiences between EM-WBI and controller-based interaction in order to generate soft design guidelines (as described in (Hornecker, 2005)) for whole body interaction.

2 Related Work

Much work in whole body interaction design has focused on explorations and the generation of soft guidelines for collaborative environments (e.g. (Bryan-Kinns, 2010; Hornecker, 2005; Hornecker and Buur, 2006; Price et al., 2009)). While some of this literature mentions embodied metaphor theory, none of it uses this theory as a means to explicitly design interactional mappings. Shoemaker et al. present several body-centric tools for interaction with large wall displays (Shoemaker et al., 2010). One is loosely based on the metaphor of storing individual user data in the user's shadow stomach. However, this form of metaphor is more rightly called a metaphorical blend (Imaz and Benyon, 2007) and is not based on an image schema.

In a qualitative study, Hashagen et al. compared whole body interaction to a desktop version of a system teaching children about the abstract domain of swarm behavior (Hashagen et al., 2009). Their interaction model maps movement speeds to virtual object speeds (identity function) and to colours (via learned conventions such as faster = red). They did not use image schema or metaphor to structure the mapping. They found that children were more able to understand the rules of swarm behavior and transfer their learning to new situations with the whole body-based system. They suggest a strong motivational factor contributed to this result.

Hurtienne (2009) has suggested and investigated the use of image schemas in the design of interaction with web and software based systems. Several researchers have suggested that image schemas and metaphor may be suitable structures to utilize for interaction with tangible or whole body interaction systems (Antle et al., 2007; Holland, 2010; Svanaes and Verplank, 2000). For example, Holland et al. studied the utility of image schemas, conceptual metaphors and blends for the design of a whole body system for learning about tonal harmony (Holland, 2010; Holland et al., 2009). Preliminary results from their comparison of whole body and desktop modes of interaction suggest that participants could very quickly learn the basic movements required to engage with the whole body system based on image schematic patterns. However, they used many image schemas, metaphors and blends in their system design which makes it difficult to disambiguate causes and effects for any one mechanism. In addition, because tonal harmony is a complex phenomenon, their system included many input and output factors, which make it an effective learning system but limits the strength of claims that can be made with it as a research instrument.

Antle et al. (2009a) have built both whole body and tangible interaction-based percussive sound production systems using single and multiple image schemas and related conceptual metaphors. These systems enable users to control simple sound concepts, such as volume, pitch and tempo, by using either whole body movements or by moving tangible objects. Studies with both children and adults indicated that participants performed better using the EM-WBI system compared to an equivalent but non-metaphor-based WBI system (Antle et al., 2009c). A study with children and tangible sound making objects produced similar findings and indicated that children were able to produce simple sound sequences using embodied metaphor-based tangible objects (Bakker et al., 2011). In these cases, embodied metaphor was used to design the interaction with perceptual constructs (e.g. percussive sound volume, tempo and pitch).

What remains is to design a system in which a single image schema and related metaphor is used to design the interactional mappings between input actions and display changes in an abstract application domain, and then compare this design to a traditional interaction paradigm. This study addresses this gap in knowledge through a design oriented empirical research study.

3 Theoretical Knowledge

We provide the theoretical foundation for our work, taking sections 3.1 and 3.2 from (Antle et al., 2011).

3.1 Embodied Conceptual Metaphor Theory

Johnson (1987) claims that conceptual metaphors arise unconsciously from experiential gestalts relating to the body's movements, orientation in space, and its interaction with objects. He calls these fundamental gestalts embodied schemas, also called *image schemas*. A metaphor involves understanding one thing in terms of another through the metaphorical relation between a target domain and a source domain. A conceptual metaphor involves understanding a concept in terms of an image schema. An image schema in the source domain is used to structure understanding of a concept in the target domain through metaphorical elaboration. Johnson suggests that a cornerstone of human meaning-making is our ability to form conceptual metaphors by using the structural and inferential properties of image schemas to structure and organize abstract concepts.

There are only a few studies that apply embodied metaphor theory in human computer interaction. The general premise of this work is that interfaces or interaction models that are consistent with metaphorical elaborations of image schemas will be more effective, efficient and satisfying to use. For a general discussion of the role of image schemas and conceptual metaphors in user interfaces see (Hurtienne et al., 2008), and in interaction models see (Antle et al., 2009a).

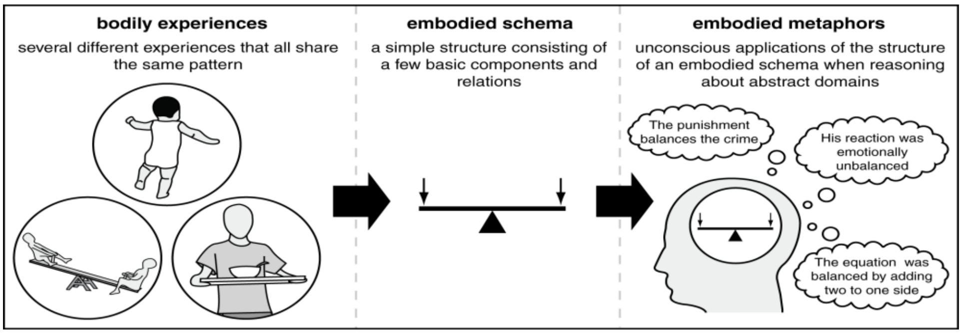
3.2 The Meanings of Balance

Johnson (1987) presents an analysis of the meaning of balance as both an experience and a concept. He states that our experience of balance is so pervasive and basic that we are seldom aware of its existence. He goes on to explain that the structure of balance is a key element that pulls our physical experience together as a coherent whole. For example, as bipeds, a toddler learning to walk immediately experiences various states of bodily balance and imbalance (Figure 1, left). We learn about balance with our bodies. Thus the meaning of balance emerges through acts of balancing our bodies. Long before we have grasped the meaning of the word we develop several image schemas for balance based on our experiences.

As children develop, image schemas related to balance begin to structure and give coherent meaning to their perceptions. In the realm of visual perception, we soon learn to interpret visual imagery as balanced or imbalanced. An image with a black circle placed at the interior edge of a square is interpreted as less balanced than an image where the circle is in the middle of the square. The image schema for bodily balance (around a point) structures this interpretation. Balance or imbalance does not objectively exist in the images. Balance comes through our act of perception and our interpretation, which utilizes a balance image schema.

Balance image schemas are also used to give meaning to balance in abstract domains such as psychological states, legal systems, mathematics, and social justice (Figure 1, right). Through metaphorical elaboration, we interpret an abstract concept of balance based on its similarities with one or more image schemas for balance. For example, when we speak of social justice, we infer that justice involves a balance of factors such as rights, privileges, damages, and duties. Our understanding and judgments arise from the twin-pan[1] balance schema (Figure 1, centre). We treat factors metaphorically like forces or weights in the pans of a scale. The scale can be imbalanced by either side of the fulcrum having too much or too little metaphorical weight or force. For a detailed discussion of balance schemas and their metaphorical elaborations, see (Johnson, 1987, Chapter 4).

Figure 1 Image schema for balance is used to structure and organize abstract concepts of balance



3.3 Definitions

One useful distinction from cognitive linguistics is between body-balance and meaning-balance. We can look for evidence of body-balance through participants' enactment (actions and movements) that takes the form of the balance image schema. For example, they might be observed moving back and forth from one leg to the other, standing still and leaning from side to side, or standing in the centre of a board to balance it. We can look for evidence of meaning-balance through participants' verbalizations of their interpretation of the balance metaphor. For example, they might talk about how pictures of shelters looked balanced because they both provide shelter for about the same number of people and are neither too elaborate nor too impoverished. In this study we look for evidence of both body-balance and meaning-balance. We would expect to see evidence of body-balance in the EM-WBI group and meaning-balance in both the EM-WBI and controller group.

4 Springboard System Design and Implementation

We describe the design and implementation of Springboard, taking excerpts about the EM-WBI implementation from a another paper (Antle et al. 2011).

[1] We use the term twin-pan to be consistent with (Johnson, 1987). Synonymous terms include: teeter-totter, seesaw and scales.

4.1 Why Social Justice?

Although evidence in support of embodied metaphor theory is not uncontested, we use it pragmatically to inform design and then look for benefit in this context. We are interested in how image schemas may be used to structure abstract concepts; how this mechanism can be leveraged in whole body interaction design (Antle et al. 2009a); and tangible interaction design (Bakker at al., 2011); and if there is a benefit to doing so. In earlier projects we explored the benefits of an embodied metaphorical interaction model in audio environments (Antle et al., 2009c). However, changes in sound parameters are largely perceptual rather than conceptual and are more physical than abstract. Therefore, in this project, we investigate the benefit of our embodied metaphor approach to whole body interaction in a more abstract conceptual domain. We decided to explore the twin-pan balance schema and the abstract concept of balance in social justice. We chose balance in social justice because it is documented in detail in (Johnson, 1987) and it is a very abstract concept. The topic also lends itself well to a large visual and sonic multimedia interactive environment. As such, it was chosen as a suitable abstract domain for this stage of our investigations.

We chose to focus on three different issues in social justice so that we could create three sets of content for our user study. Based on pilot studies, we chose issues related to food production, shelter production and community safety. For brevity, we abbreviate these as: food, shelter and safety. Having three content sets allowed us to separate effects due to interaction design strategies from effects due to choice of content. For consistency of experience, each issue was represented in the Springboard system using the same interaction and display designs. Each issue was also conceptualized based on metaphorical extension of the twin-pan balance schema (Figure 1, centre) as described below in section 4.6, Image Display.

4.2 Design Goals

Our main objective was to create a system that we could use as a research instrument to address our research questions by comparing the similarities and differences between EM-WBI and controller-based interaction for the same multimedia environment.

4.2.1 EM-WBI and Controller-based Interaction

One design goal was to create a system that supported both an EM-WBI and controller-based mode of interaction with the same content and display system. The interaction mode is the mapping layer that relates body movement-based input or controller-based input to changes in displayed images and sounds. For comparative purposes, we decided to use three EM-WBI modes and use two controllers: a rotational dial and a vertical slider.

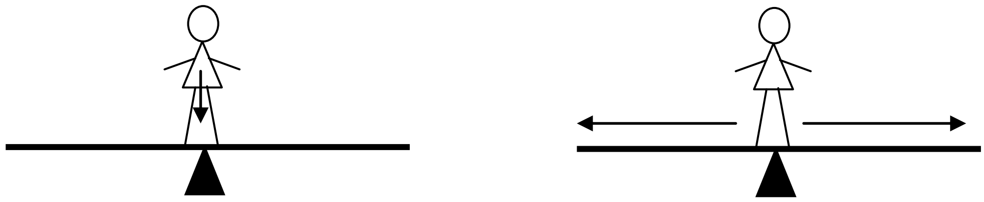
We discuss the goals for the design of three EM-WBI modes in 4.2.2 and the goals for the design of the two controller-based modes in 4.2.3.

4.2.2 Body-centric and Spatial Enactments of Twin-Pan Balance Image Schema

To continue to investigate the differences between sensing spatial versus body-centric enactments of a image schema (Antle et al., 2011), another design goal was to develop three EM-WBI modes, described here. The twin-pan balance schema arises from both our experiences of balancing our body in a specific space and from bodily balance (e.g.

walking). For example, when a biped stands on a teeter-totter or seesaw (Figure 2), they can feel in or out of balance based on where their centre of gravity is (body-centric experience giving rise to the schema), or based on their spatial position on the teeter-totter (spatial experience giving rise to the schema), or both. Based on earlier work in an audio environment, Sound Maker (Antle et al., 2009c), we found that participants tended to give priority to the spatial rather than body-centric enactments and interpretations of the schemas in the system, although we observed that both operated together. Conversely, in our initial qualitative study of Springboard, we found few differences between users' experiences with a system implemented using the body-centric versus the spatial image schema (Antle et al., 2011). We needed to explore this issue further. Therefore, we wanted to have three whole body interaction modes: one based on sensing only spatial position in a defined input space (spatial) (Figure 2, right), one based on sensing only centre of gravity (body) (Figure 2, left), and one in which both are sensed and amalgamated (body + spatial) (Figure 2, both).

Figure 2 Twin-pan balance schema: body-centric, balanced centre of gravity (left), and spatial, balanced position in space (right)



4.2.3 Same Display System but Different Interaction Modes

A third design goal was that while the interaction mode could be varied, the sound and image content sets and the display engines would remain the same for all five interaction modes (spatial, body, spatial + body, dial, slider). Image and sound content were to be structured using *meaning-balance*. The visual layout of images on the screen was to be consistent with the twin-pan two factor structure (i.e. side by side). However, implementation of the input space would be different for each of the five modes.

For the EM-WBI modes, our goal was to use the twin-pan balance schema to structure the input space and to map changes in input data to metaphorically consistent changes in the display of images and sounds. For the controller-based interaction modes our goal was to not to use the balance image schema to structure the input space. Instead, our goal was to use the dial controller to produce rotational position data which would be used to control the images and sounds. Similarly, our goal was to use the vertical slider controller to produce vertical position data which would be used to control the images and sounds. However, for these modes, the balance schema was to be used to structure the content and visual layout exactly in the same way as in the EM-WBI modes. The display system would not change, only the input style would change. For example, turning the dial all the way to one endpoint should produce the same display changes as moving the body completely out of balance. See sections 4.4 - 4.8 for details of system implementation.

4.2.4 Balancing Enactment and Perception for EM-WBI modes

A fourth design goal was that the whole body version of the environment should support the user to both move in the input space and perceive changes in the display without privileging one more than the other. For example, input movements should not be too difficult or to trivial to enact. Similarly, changes in output images and sounds should be fairly easy to perceive (while moving).

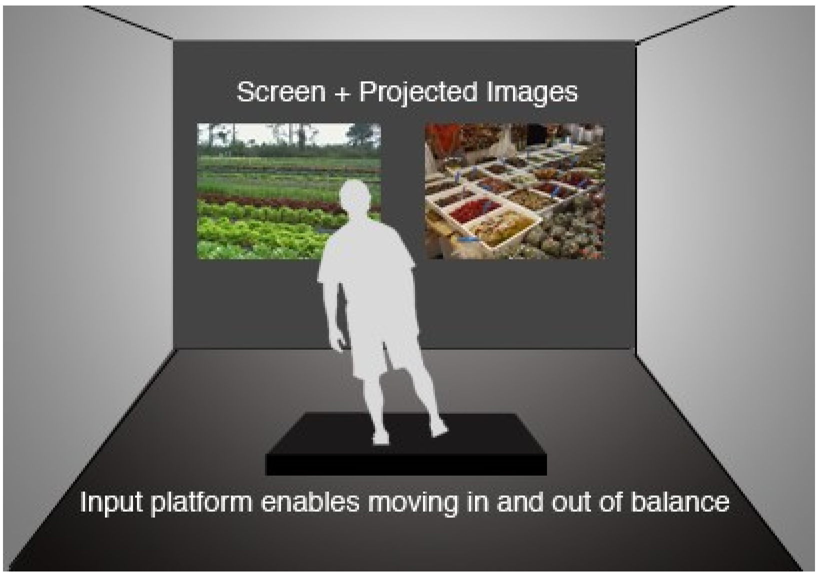
4.3 Interpretations of Balance: Values and Subjectivity

Concepts of balance in social justice are value laden and subjective. We have chosen issues and related themes that can be conceptualized along a continuum using the twin-pan balance schema. We designated the position on the balance spectrum for each image through a collaborative triangulation process involving image sorting by three researchers working on the project. Our design objective was to support users to explore these issues visually and sonically in order to investigate usability, cognitive-motor and experiential factors through a comparative study. It is not necessary that users agree with us about what constitutes a balanced solution to a particular issue in order to investigate these factors.

4.4 The Springboard Multimedia Interactive Environment

Springboard was developed to enable our investigation comparing whole body interaction to controller-based interaction. The Springboard environment supports users to interact through whole body interaction or controller-based interaction to explore images (Figure 3) and sounds related to three social justice issues.

Figure 3 Springboard system image display space



For the whole body interaction modes (spatial, body, body + spatial), the active input space is a small raised platform (132 x 71 x 20 cm) made from a crib mattress spring, board and black cloth (Figure 4, left). Since standing in a balanced way is a normal state for most adults we required an input space that upset this balance but not so much as to focus the user away from the display space. When a user steps onto the platform, their centre of gravity immediately becomes slightly out of balance since they will likely wobble on the platform (Figure 4, right). The rectangular design of the platform also supports lateral movement. By moving left or right, the user can also be out of balance spatially and the design of the platform ensures that it is even more wobbly at the edges than in the centre. States of bodily balance are determined as users move their body's centre of gravity and spatial position on the spring board input platform. Using a camera vision system, we track centre of gravity compared to foot position to establish body balance (Figure 7) and track spatial position on the platform to establish spatial balance. This allows us to have three whole body interaction modes: spatial position only (spatial), body centre of gravity only (body), or an aggregate of body and position (body + spatial). For the two controller modes, dial and slider, the user sits in a canvas chair and places the controller on their lap. The same device was used to implement the two controller modes (Figure 5).

Figure 4 Input platform (left) and input space (right)

Figure 5 Controller device with rotational dial (top) and vertical slider (bottom)

4.5 Summary of Interaction Modes

The Springboard system can be configured with five different interaction modes: three configurations involve EM-WBI (spatial, body, body + spatial) and two configurations involve controllers (dial, slider). The controller based mappings do not utilize image schema in their mappings since the dial is rotational and slider is up-down, neither of which correspond to the twin-pan balance schema. We considered designing a tangible twin-balance object which could be used to control the system, but we will leave that for our next study. In this study we use traditional controller operators.

All of these interaction modes can be used to interact with the system by changing the displayed images and sounds as described below.

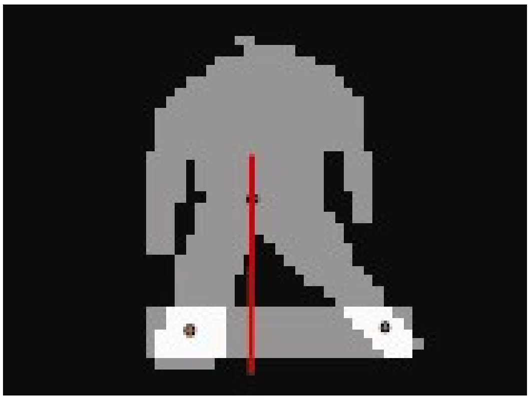
4.6 Image Display

Pairs of images are displayed on a large wall screen (Figure 3). Each pair of images depicts some degree of balance or imbalance in terms of a social justice issue. For example, in Figure 6 the images are related to the issue of equitable use of resources for shelter for all people. One image shows an expensive, resource-intensive living room. The other image shows a person using a cardboard box for shelter. These two images show extreme imbalance in shelter resource use. In the whole body modes, these images would be displayed if the user was out of balance bodily, spatially, or both, depending on the interaction mode. Using the controllers, these images would be shown at one endpoint of the vertical slider or dial rotation. The pairs of images synchronously fade in and out as the user's input actions change. For the EM-WBI modes, the user's movements in and out of balance trigger metaphorically related changes in the images depicting the balance of two factors for each social justice issue. For the controller modes, the user rotates or slides the dial to traverse the display space.

A description of the image display engine follows. Each of the food, shelter and safety issues had two factors that can be balanced or imbalanced to various degrees, as described above. The factors are depicted with pairs of images displayed on a large vertical screen as shown in Figure 3. For each issue, a set of images depicting different aspects of that issue were sourced and then tagged through a collaborative sorting process in order to categorize them in five numbered bins. For each issue, the bins range from (1) too much of the factor, to (5) too little of the factor with a central bin (3) for balanced factors. For example, for the shelter issue, an image of opulent interior of a private residence was tagged 5 (too much quality/resource consumption) and an image of a person sleeping in a cardboard box was tagged 1 (too little quality/resource consumption), as shown in Figure 6. Each of the bins contains many images to support variation and multiple interpretations.

Figure 6 Display layout for shelter (out of balance)

Figure 7 Body centre of gravity balance index



4.7 Sound Display

The sound feedback for Springboard utilizes several approaches to representing the concept of balance through sonic aspects. The sound design is described in detail in (Droumeva et al., 2009). We summarize the sound design here. The sound feedback provides constant ambient information, structured using the twin-pan balance schema, which responds to and guides user actions. The obvious choice of a left-right channel panning was discarded. Panning is a representation of balance based on a cultural invention associated specifically with the technology of headphones. It is not based on a metaphorical extension of bodily or perceptual sonic balance. In addition, panning does not provide a clear resolution of sound change. We focus on more primary perceptually-based sound parameter changes such as pitch, timbre and phase, in order to achieve a sense of sonic balance and imbalance.

4.8 Camera Vision System

For the whole body modes, the bodily balance of the user is determined using a blob tracking and analysis computer vision system developed in the Max/Jitter programming environment. The participant stands in front of a black background on the black platform (Figure 4, right). This setup allows a simple background subtract process to be used to isolate the participant's image. The total balance of the participant is calculated using a body centre of gravity balance index (Figure 7) and a spatial balance index. Depending on the mode, the two indices are each used individually, or are combined using a scaled addition process, producing the total balance index that ranges from -10 to +10 where 0 reflects a completely balanced body state. The total balance index is used to control the image and sound display engines.

4.9 Limitations of Springboard as a Research Instrument

Designing an interaction model and implementing a sensing and reasoning system based on an embodied metaphor is difficult and relies on several simplifying decisions. The choice to use the twin-pan balance schema was driven by the focus on justice in which the scales (twin-pan) are a dominant concept (Johnson, 1987). The choice of sensing user's centre of gravity and spatial location on the platform in order to determine states of bodily balance was largely driven by Johnson's (1987) work, body-storming exercises (Oulasvirta et al., 2003) and exploration of different structures that would cause users to move in balanced and unbalanced ways. The classification of images into five bins (rather than, say, three or seven) was chosen to ensure that the environment was interesting and yet not difficult to understand and interpret. The decision to have two images depicted side by side on a large screen was made to be consistent with the two factors or twin-pans. All of these decisions have implications for the use and interpretation of the environment and the strength of knowledge claims made from user studies. However, the five system modes were identical except for the way that users interacted with the system, which supports using Springboard as a research instrument to compare the effects of whole body and controller-based interaction on usability, motor-cognitive processes and user experience.

5 User Study Methodology

A comparative user study using a between groups design was created to address the research questions posed above. In this paper, we focus on the quantitative system log and survey-based measures in order to address our research questions. We investigate how EM-WBI compares to controller-based interaction in terms of enabling users to explore Springboard from a perspective of usability (efficiency, effectiveness, extent of exploration, satisfaction with performance), cognitive-motor qualities (awareness of schema, ability to focus on content) and felt experience (enjoyment, interest, competence, impact).

5.1 Participants

The study was comprised of sessions with 76 adult volunteers of both genders (40 male, 36 female), predominantly aged 18–25 years old (96%). Volunteers were recruited from an urban university campus. All participants used computers daily. Only 11% were experienced users of interactive environments, however another 52% had used interactive environments one or more times.

5.2 Study Design

Seventy six participants were randomly assigned to one of five groups based on each interaction mode for Springboard. Groups were: spatial (n=15), body (n=14), spatial + body (n=16), slider (n=15) and dial (n=16). The groups were roughly gender balanced. This study is not a controlled experiment since many factors vary between the EM-WBI and controller interaction groups. Yet we can still discuss results in terms of groups and make statistical comparisons using data from each group to look for similarities and differences between usability, motor-cognitive and experiential measures. Strong claims cannot be made from this study, however, we can explore the similarities and differences and generate considerations which can be used to inform future designs (what is real) rather than provide evidence for theories (what is true) (Fallman, 2007).

5.3 Study Design Limitations

Our research design does not separate out the effects of whole body interaction from embodied metaphor-based interaction. That is, we do not have a condition with whole body interaction but no embodied metaphor. Based on our previous work (Antle et al., 2009c), we suggest that using embodied metaphor to structure whole body interaction is an effective strategy which will highlight some of the advantages of whole body interaction. Using arbitrary mappings is possible, but then the mappings must be learned or discovered rather than enacted "intuitively" as discussed in Antle et al. (2009b). In addition, in whole body environments without tangible objects, there is often little physical form to offer affordances for action (Bakker et al., 2011). This means that participants may take much longer to learn how particular actions produce desired system responses. When interacting with a controller, such as a slider or dial, the physical form constrains possible actions to linear motion and rotational motion with set endpoints and the mapping is easily learned and enacted. Thus an EM-WBI model provides an effective experience which is commensurate with a simple controller.

5.4 Procedure

Seventy six adult users completed a training session and three tasks. Each task involved having the participant use Springboard to explore a multimedia content set related to food, shelter and safety issues. Users were told to explore the images and sounds, and to stop moving (their body or the controller) and verbally indicate when the images and sounds represented "balance" for that issue. The tasks were followed by a series of qualitative interview questions, which are reported on in Antle et al. (2011), and a set of written survey questions. The survey forms the basis for the results presented in this paper. We also conducted post-survey phenomenological style interviews with select participants but have not yet reported on this data.

5.5 Measures

For each task (i.e. for food, shelter and safety), we collected data related to each construct identified in the eight research questions (abbreviated RQ). Data included system log data and a post-session survey. The survey included structured questions using various quantitative scales. For usability measures, we logged task time (RQ 1a. Efficiency), logged which images (bin numbers) were being displayed when the user indicated that the issue was balanced (RQ 1b. Effectiveness), and recorded how extensive their exploration of the content set was based on frequency of bins sampled (RQ 2. Extent of exploration i.e. how much of the content set they viewed). We asked users to rate their satisfaction with their performance using a 1 to 7 Likert scale (RQ 3. Performance satisfaction).

For the motor-cognitive qualities, we included three questions in the post-survey. Two questions asked about their awareness of how they interacted with the system (RQ 4. Awareness). In the first question they were given a set of pictorial representations of various image schemas to choose from and asked to circle the one that best represented how they purposely interacted with the system. We included schemas for twin-pan balance, up-down, in-out, and linear path. The second question was the same but the choices were given in words (e.g. up-down, in-out of balance). We also asked them to rate their ease of focusing on the image content while using the system (RQ 5. Focus).

For the experiential constructs we used two subscales of a validated survey called the Intrinsic Motivation Inventory (IMI) (Ryan, 2006) to collect data about enjoyment and interest (RQ 6. Enjoyable), and competence (RQ 7. Competency). For the impact construct we asked participants two pre and post questions (RQ 8. Impact). In the first question, we asked them to rate their awareness of issues related to food, shelter and safety in social justice before and after the session. Then we compared the impact of their experience in terms of a change in their awareness of social justice issues. We also asked participants a pre and post question asking them to rate their willingness to take action about each issue before and after their session. While we expected there to be a positive effect based on their experience, we considered the magnitude of the pre to post change for both of the impact questions across groups.

Data analysis was done using parametric tests when interval or ordinal data was normally distributed with equal variances, and non-parametric tests when data was binary or had unequal variances. The one way ANOVA (parametric) and Kruskal-Wallis tests (non-parametric) were used to compare means between the five groups. Where no significant differences existed between the three EM-WBI groups, or between the two controller groups, we collapsed data into two groups. In this case, t-tests (parametric) and the Mann-Whitney test (non-parametric) were used to compare the aggregated data for the EM-WBI and controller groups. Repeated measures ANOVA was used to compare pre and post test scores.

6 Results

We present the results of our comparative study using a dial and slider controller as the baselines compared to three modes of EM-WBI. We report findings for the three EM-WBI groups and the two controller groups as "whole body" and "controller" when data has been collapsed.

6.1 RQ 1a. Efficiency: Task Times

There were no significant differences between the time taken to complete each task between the whole body and controller groups for any of the three tasks. However, there were significant differences between the times for each task. The first task took on average, 161 seconds for the whole body group and 148 seconds for the controller group to complete (Table 1). The second and third tasks took, respectively 84 and 86 seconds for the whole body group and 95 and 114 seconds for the controller group. There are large variations in all the task times (Table 1). Results from a one way ANOVA for both whole body and controller groups indicated that the average task time was significantly different between tasks at the $p < .0001$ level for whole body and controller groups ($F(3, 166) = 6.53$, $p < .0001$ and $F(2,117) = 14.83$, $p < .0001$). Results from Tukey post hocs indicated that users took significantly longer for the first task than for the second and third tasks for both groups. This result reflects a learning curve with the system. It is interesting to see that the whole body group accomplished the second and third tasks slightly faster on average, although not significantly so, than the controller group. We suggest two interpretations. Either the participants in the whole body group were more able to quickly "find balance" using their bodies or participants in the controller group spent longer looking at images (e.g. enjoyed flipping through images in a leisurely way). The structure imposed by giving participants a goal-related task and our observations that participants were very task focused support the first interpretation.

	Task 1: Food		Task 2: Shelter		Task 3: Safety	
	Mean (s)	Std Dev (s)	Mean (s)	Std Dev(s)	Mean (s)	Std Dev (s)
WB	161	155	84	63	86	60
Controllers	148	68	95	69	114	94

Table 1 Mean task times and standard deviations (seconds)

6.2 RQ 1b. Effectiveness: Accuracy

Participants were asked to indicate when the images on the display depicted balance in social justice. There were no significant differences between the participants' ability to complete each task by stopping their interactions with their body or a controller when the image pair they thought depicted balance was displayed.

6.3 RQ 2. Extent of Content Exploration

The amount of variation in exploration of the image sets was significantly greater for the dial group than the slider or whole body group at the $p < 0.05$ level ($F(4,72) = 3.48$, $p = .013$). Dial users explored the whole range of pictures more so than the slider or whole body group. In part, we suggest this is because the dial was very easy to turn all the way to each endpoint with minimal effort. However, to reach the extremes of whole body balance required participants to assume an unbalanced position, which isn't very comfortable, nor can these kinds of positions be held for any length of time. We also noticed that small or shorter participants had more difficulty reaching the extreme unbalanced states. This is due to limitations of the blob sensing and interpretation algorithm and should be addressed in future system iterations.

6.4 RQ 3. Performance Satisfaction

The Likert scale-based response data for the two performance questions had equal variances so t-tests were used. T-tests indicated that the controller group was significantly more satisfied with their performance at the $p < 0.01$ level ($F(74) = .008$, $p = .010$) and felt that they were pretty good using the system at the $p < 0.05$ level ($F(74) = .269$, $p = .015$) compared to the whole body group.

6.5 RQ 4. Awareness of Twin-Pan Balance Schema in Interaction

The multiple choice questions for schema analysis were coded 1 for correct and 0 for incorrect. Binary data was analyzed using Mann Whitney for independent samples, with the data aggregated into two groups because there were no significant differences between the three whole body groups or the two controller groups. When participants were asked how they interacted with the system and were given a set of pictorial representations of various image schemas to choose from, participants in the whole body group chose the twin-pan balance schema significantly more often than those in the controller group at the $p < .01$ level ($U = 512$, $p = .004$).

When we asked the same question but provided word-based descriptions of the schemas, participants in the spatial whole body group were significantly more able to choose the balance schema ("in and out of balance") than all the other groups at the $p < .05$ level ($U = 511$, $p = .023$).

Frequency analysis showed that participants in the spatial group picked the balance schema using pictorial representation 93% of the time and using words 87% of the time compared to 79% (picture) and 57% (words) for the body group and 88% (picture) and 56% (words) for the spatial + body group. This is consistent with our earlier findings in which we found that participants interpret a whole body environment primarily using spatial schemas to structure exploratory behaviors (rather than body-based schemas), and that participants often used a spatial schema to interpret how the system worked (Antle et al., 2009c).

6.6 RQ 5. Ability to Focus on Multimedia Content

T-test results indicated that participants using a controller found it significantly easier at the $p < .05$ level to focus on the images than those using whole body movements to control the system ($F(74) = 1.027$, $P = .032$). There were no significant differences between participants' ratings of their ability to focus on sounds between groups.

6.7 RQ 6. Enjoyment and Interest

There were no significant differences in participants' ratings of their enjoyment or interest in using Springboard between any groups based on the IMI subscale for enjoyment and interest. We have found a lack of significant difference using Likert scales in past work where participants voiced qualitative differences in felt experiences. It is unclear whether a seven point Likert scale is a sensitive enough research instrument to capture these kinds of differences. For example, participants in each group may enjoy the experience for different reasons. We leave this to future methodological research.

6.8 RQ 7. Competence

There were no significant differences in participants' ratings of their feelings of competence using Springboard between any groups based on the IMI subscale for competence.

6.9 RQ 8a. Impact: Awareness of Issues in Social Justice

Using Springboard had a positive impact on participants in terms of their ratings of self-awareness of issues in social justice. Using paired t-tests showed that pre to post-session ratings increased across all groups very significantly at the $p < .0001$ level ($t = -4.99$, $p < .0001$). This means that Springboard sessions impacted all participants in terms of increasing their awareness of issues in social justice, as expected. Correlation analysis indicated that pre and post test ratings for participants were significantly correlated at the $p < .0001$ level ($r = .581$). This means that participants' pre and post ratings were related which adds validity to our results. Repeated measures indicated that whole body group ratings increased significantly more than the controller group at the $p < .0001$ level ($F(1) = 741.6$, $p < .0001$). While both groups were impacted by their experiences, the whole body group was more affected by or felt more impact from their experience than those sitting in the deck chair using a controller.

6.10 RQ 8b. Impact: Willingness to Take in Action Related to Achieving Social Justice

Using Springboard had a positive impact on participants in terms of their ratings of their willingness to take action related to achieving social justice. Using paired t-tests showed that pre to post session ratings increased across all groups very significantly at the $p < .001$ level ($t = -3.34$, $p < .001$). This means that Springboard sessions impacted all participants, also as expected. Correlation analysis indicated that pre and post test ratings for participants were significantly correlated at the $p < .0001$ level ($r = .857$). Repeated measures indicated that whole body group ratings increased significantly more than the controller group at the $p < .0001$ level ($F(1) = 609.0$, $p < .0001$). While both groups were impacted by their experiences, the whole body group was more affected in terms of their willingness to take action towards achieving social justice than those sitting in the deck chair using a controller.

7 Discussion

Based on our earlier work, we expected that including embodied metaphors in a whole body based interaction model would have usability and other benefits over a simple control device due to the schematic structure of the interaction model. However, our findings did not indicate usability or satisfaction benefits. This is counter to suggestions by Hurtienne et al. (2008) but most likely an effect caused by the whole body nature of interaction rather than the embodied metaphor. A follow-up study comparing embodied metaphor in both whole body and controller mappings could disambiguate these effects.

Our results provide important contributions to research involving the design of whole body interaction for abstract (rather than spatial or concrete) domain applications. First, we found evidence that whole body interaction with an abstract domain results in a more *impactful* experience with the content than using controller-based interaction. We might expect this benefit with concrete domains; for example, through learn-by-doing style interaction as suggested by Klemmer et al. (2006). However, our study is the first systematic study to provide empirical evidence of this effect in an abstract domain. Our findings are in line with Holland et al. (2010) who report anecdotal observations about benefit in a case study of the Haptic Drumkit. It is also important to consider that we found that EM-WBI supports a greater *impact* of felt experience at no time or accuracy *usability* cost. These findings may be useful for designers of whole body interaction applications in domains including: learning, training, simulation, art installations, museums or other public exhibits and gaming.

A second and more important contribution of our work is that we ground our predictions and study on the specific motor-cognitive mechanism of image schemas and related metaphors. This mechanism can be explicitly supported through specific interaction design methods (as described in (Antle et al., 2009a; Bakker et al., 2011)). The derivation of embodied interaction models based on embodied metaphor theory moves current research in the field past descriptive accounts of the benefits of whole body interaction into the realm of scientific explanations and testable mechanisms. It also provides grounding for general design recommendations across a range of application areas and platforms that utilize whole body interaction.

7.1 Soft Recommendations

Our findings suggest that if the goal of an interactive environment is to facilitate quick and easy access to multimedia images and sounds, then a simple controller device is efficient and effective. The controller approach also supports participants in focusing on imagery rather than having to split their attention between the images and using their body as an input device. Participants will be better able to perform with or control the system with the controller than a system using the body as an input device. However, we suggest that there may be times when focused awareness on action and thought may create an opening for a deeper experience. This is somewhat analogous to learning a technical sport, such as golf or tennis, when a person must focus on conceptual understanding in conjunction with developing physical competence to achieve mastery. In abstract domains, deep understanding may too require mastery through effortful action and attention.

Using EM-WBI had several other positive effects that are worth considering. First, participants in this group were aware of the balance schema and may have used it to influence their interpretation of the content of Springboard. In this way, designers can enable specific image-schematic interactions and metaphorical interpretations. Second, participants who engaged with the system using their bodies felt more aware of issues in social justice; they were more affected or influenced by their experience than those who simply used a controller to flip through the image pairs. Users in the whole body group also stated that they were more likely to take concrete actions to positively contribute to social justice issues. We suggest that designers consider whole body interaction when designing to support deep and rich engagement with the content of an application. And third, since there were no significant differences in time, accuracy, enjoyment, interest or competence between the two interaction styles, we suggest that designers may want to use metaphor-based whole body interaction as a design approach which may have greater impact with little usability cost.

We envision that these recommendations will be useful for art installations, museum exhibits, public displays for engagement with social issues, learning environments, and other interactive environments where whole body interaction is possible and a design goal is deep, rich experience.

8 Conclusions

We report on a quantitative comparative study of the benefits of EM-WBI compared to control-based interaction. There were few usability or user experience costs to the whole body interaction approach; it was comparable with simple controllers. There were experiential benefits to EM-WBI approach. Users in the EM-WBI groups had a more impacted felt experience of the social justice issues.

One of the main contributions of this work is the validation of embodied metaphor theory in the domain of HCI. A second contribution is the demonstration and validation of the approach of using embodied metaphors to design interactional models for whole body systems in abstract rather than concrete or spatial domains. Lastly, based on our comparison of EM-WBI and controller-based interaction, we provide guidelines for such design work, which we envision will be useful to designers of a range of interactive technologies that support a wider range of input actions than traditional desktop configurations.

To clarify the factors at work, we suggest that a future study should compare WBI and controller designs based on embodied metaphor with non-metaphorical approaches.

Acknowledgements

We gratefully acknowledge the financial support of a NSERC Discovery Grant for personnel costs associated with running the user study; a NSERC RTI grant for equipment purchases; and a SSHRC RDI grant to support methodological work related to this study. We would also like to thank Milena Droumeva for her early conceptual work on the project and her later work on sound design; Josh Tanenbaum for his work sourcing imagery; and Sijie Wang for his support during the user studies. We appreciate the thoughtful comments and suggestions for improvements from the reviewers.

References

Antle, A.N. (2007) 'The CTI framework: Informing the design of tangible systems for children', Paper presented in the Proceedings of the *Conference on Tangible and Embedded Interaction*, ACM Press, Baton Rouge, LA, pp. 195-202.

Antle, A.N., Corness, G., Bakker, S., Droumeva, M., van den Hoven, E. and Bevans, A. (2009a) 'Designing to support reasoned imagination through embodied metaphor', Paper presented in the Proceedings of the *Conference on Creativity and Cognition*, ACM Press, Berkeley, CA, USA, pp. 275-284.

Antle, A.N., Corness, G. and Droumeva, M. (2009b) 'Human-Computer-Intuition? Exploring the cognitive basis for intuition in embodied interaction', *International Journal of Art and Technology*, Vol. 2, No. 3, pp. 235-254.

Antle, A.N., Corness, G. and Droumeva, M. (2009c) 'What the body knows: Exploring the benefits of embodied metaphors in hybrid physical digital environments', *Interacting with Computers: Special Issue on Enactive Interfaces*, Vol. 21, No. 1-2, pp. 66-75.

Antle, A.N., Corness, G. and Bevans, A. (2011) 'Springboard: Designing Image Schema Based Embodied Interaction for an Abstract Domain', in England, D. (Ed.): *Human-Computer Interaction Series*: *Whole Body Interaction* (2011), Springer, 7-18.

Bakker, S., Antle, A.N. and Van den Hoven, E. (2012) 'Embodied metaphors in tangible interaction design', *Personal and Ubiquitous Computing*, 16, 433-449.

Bryan-Kinns, N. (2010) 'Collective Body Engagement', Paper presented at the *ACM CHI 2010: Whole Body Interaction Workshop*. Available at http://lister.cms.livjm.ac.uk/homepage/staff /cmsdengl/WBI2010/ (accessed June 7 2012).

Droumeva, M., Antle, A.N., Corness, G. and Bevans, A. (2009) 'Springboard: Exploring embodied metaphor in the design of sound feedback for physical responsive environments', Paper presented in the Proceedings of the *International Conference on Auditory Display*, http://icad.org/node/3059 (accessed June 7 2012).

Fallman, D. (2007) 'Why research-oriented design isn't design-oriented research: On the tensions between design and research in an implicit design discipline', *Knowledge, Technology, and Policy*, Vol. 20, No. 3, pp. 193-200.

Hashagen, A., Büching, C. and Schelhowe, H. (2009) 'Learning abstract concepts through bodily engagement: a comparative, qualitative study', Paper presented in the Proceedings of the *Conference on Interaction Design and Children,* ACM Press, Como, Italy, pp. 234-237.

Holland, S. (2010) 'Asymmetrical multi-User co-operative whole body interaction in abstract domains', Paper presented at the *CHI 2010: Whole Body Interaction Workshop*, available at http://lister.cms.livjm.ac.uk/homepage/staff/cmsdengl/WBI2010/ (accessed June 7 2012).

Holland, S., Marshall, P., Jon Bird, J., Dalton, S., Morris, R., Pantidi, N., Rogers, Y. and Clark, A. (2009) 'Running up Blueberry Hill: prototyping whole body interaction in harmony space', Paper presented in the Proceedings of the *Conference on Tangible and Embedded Interaction,* ACM Press, Cambridge, UK, pp. 93-98

Hornecker, E. (2005) 'A design theme for tangible interaction: Embodied facilitation', Paper presented in the Proceedings of the *Conference on Computer-Supported Cooperative Work*, Springer, Paris, France, pp. 23-44.

Hornecker, E. and Buur, J. (2006) 'Getting a grip on tangible interaction: A framework on physical space and social interaction', Paper presented in the Proceedings of the *Conference on Human Factors in Computing Systems*, ACM Press, Montreal, Quebec, Canada, pp. 437-446.

Hurtienne, J. (2009) 'Cognition in HCI: An ongoing story', *Human Technology*, Vol. 5, No. 1, pp. 12-28.

Hurtienne, J., Weber, K. and Blessing, L. (2008) 'Prior Experience and Intuitive Use: Image Schemas in User Centred Design', in P. Langdon, P.J.C., & P.Robinson (Ed.), *Designing Inclusive Futures*, Springer, London, pp. 107-116.

Imaz, M. and Benyon, D. (2007) *Designing with Blends: Conceptual Foundations of Human Computer Interaction and Software Engineering*. MIT Press, Cambridge, MA, USA.

Johnson, M. (1987) *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. University of Chicago Press, Chicago, IL.

Klemmer, S., Hartmann, B. and Takayama, L. (2006) 'How bodies matter: Five themes for interaction design', Paper presented in the Proceedings of the *Conference on Designing Interactive Systems*, ACM Press, University Park, PA, USA, pp. 140-149.

Oulasvirta, A., Kurvinen, E. and Kankainen, T. (2003) 'Understanding contexts by being there: case studies in bodystorming ', *Personal and Ubiquitous Computing*, Vol. 7, No. 2, pp. 125-134.

Price, S., Falcao, T.P., Sheridan, J.G. and Roussos, G. (2009) 'The effect of representation location on interaction in a tangible learning environment' Paper presented in the Proceedings of the *Conference on Tangible and Embedded Interaction,* ACM Press, Cambridge, UK, pp. 85-92.

Ryan, R.M. (2006) 'The Intrinsic Motivation Inventory (IMI)', available at http://www.psych.rochester.edu/SDT/measures/IMI_description.php (accessed Feb 18 2011).

Shoemaker, G., Tsukitani, T., Kitamura, Y. and Booth , K.S. (2010) 'Whole body large wall display interfaces', Paper presented in the Proceedings of the *Conference on Human Factors in Computing Systems,* ACM Press, Atlantic, GA, USA, pp. 4809-4812.

Svanaes, D. (2001) 'Context-aware technology: A phenomenological perspective', *Human Computer Interaction*, Vol. 16, pp. 379-400.

Svanaes, D. and Verplank, W. (2000) 'In search of metaphors for tangible user interfaces', Paper presented in the Proceedings of the *Conference on Designing Augmented Reality Environments,* ACM Press, Elsinore, Denmark, pp. 121-129.