Final published version appears in Interacting with Computers 21 (2009) 66–75.
https://doi.org/10.1016/j.intcom.2008.10.005

What the body knows: Exploring the benefits of embodied metaphors in hybrid physical digital environments

Alissa Antle, Greg Corness, Milena Droumeva

School of Interactive Arts and Technology, Simon Fraser University, 250-13450 102 Avenue, Surrey, BC, Canada V3T 0A3

ABSTRACT

A recent trend in ubiquitous computing is the development of new forms of interfaces, which rely on *embodied interaction*. We focus on the definition of embodiment that refers to the ways in which abstract concepts rely on metaphorical extensions of embodied schemata shaped by processes below the level of conscious awareness as explored by Lakoff and Johnson [Lakoff, G., Johnson, M., 1980. Metaphors We Live By. Chicago Press, Chicago, IL, USA]. Our inquiry focuses on understanding the role embodied metaphors may play in supporting people to understand the possibilities for physical interaction in augmented spaces. We explore this issue through the development and evaluation of an interactive audio environment. We instantiate metaphor theory by using embodied schemata as the basis for the interactional metaphor that relates full-body input actions to audio output responses. We demonstrate and explore the benefits of this approach through a comparative experiment in which adults and children learn to use our audio environment. The results from our experiment indicated that embodied metaphors improve usability however, other factors including discoverability, perceivability of feedback and duplicity of structural isomorphism may mediate these metaphor-based benefits. We have generalized our main findings as a set of suggestions for the design of embodied style interfaces that rely on physical interaction.

1. Introduction

> "It is not enough to say that the mind is embodied; one must say how." (Edelman 1992, p. 15 in Gallagher (2005).

As computation moves off the desktop and into the environment, successful design of augmented environments requires that we understand, leverage and extend the ways in which human intelligence and experience are determined at a fundamental level by the details of human physical bodies embedded in a material world. There are several complementary views on human-world relations that seek to understand human cognition beyond the inner workings of an individual mind. An embodied view of cognition grants the body, situated in the environment, a central role in shaping the mind. The critical concept underlying embodied cognition is that humans only think the way they do because they have human bodies and live in the social, physical environment that they live in.

An embodied view of cognition is commonly introduced in contrast to the Cartesian or Cognitivist tradition. Cognitivism treats cognition as something that occurs in the mind, separate and irrespective of the body. This perspective was largely inherited by the cognitive science and human computer interaction communities of the latter half of the 20th century. This view resulted in the conceptualization of the mind as an abstract information processor that arose with the genesis of digital computation (Wilson, 2002). While this perspective has merit for the design of certain types of computation that focus on symbol manipulation (e.g., expert systems), it is not sufficient for understanding all aspects of human cognition (Clark, 1999). An embodied perspective provides an alternative view which has merit for situations where the roles of the perceptual and motor systems and their interaction with the environment are paramount (e.g., robot locomotion, learning in augmented environments) (Clark, 1997).

The philosophical origins of an embodied view of cognition can be traced to the phenomenological tradition begun by Husserl and continued by Heidegger. Phenomenology is based on a practice of reflection upon the structures of lived experience and posits a central role for perception and action in meaning-making. This theme is extended by Merleau-Ponty in his proposal that consciousness, the world, and the body are intertwined, mutually engaged and dynamically coupled (2002). It is also evident in Wittengenstein's oft repeated quote "If a lion could talk we could not understand him" (1953).

Outside of philosophy, this theme can be observed in the work of several significant theoretical thinkers in psychology and cognitive science. For example, this view resonates with the genetic epistemology of developmental psychologist Piaget who emphasized the emergence of cognitive abilities grounded in sensori-motor abilities (1952). It relates to the ecological psychology of Gibson who posits a view of perception based on affordances – potentials for interaction with the physical environment (1979). More radically, this view is endorsed by the enactive cognitive science of Varela, Thompson and Rosch who stress the interplay of the world and the structure of the knower (1991). Linguists Lakoff and Johnston also built on this theme when they explored how abstract concepts are based on metaphors of bodily, physical experiences (1980).

An embodied view of cognition was popularized in the human computer interaction community by Dourish. He used the term *embodied interaction* to describe an approach to interaction design that places an emphasis on understanding and incorporating our relationship with the world around us, both physical and social, into the design and use of interactive systems (Dourish, 2001). An embodied view on interaction provides us with an interpretive perspective that can be used to describe and explain the actions and interactions of users with ubiquitous computing applications. It promises design that recognizes the role of experience and the coupling of action and meaning (Imaz and Benyon, 2007). However, to date, there has been more work that deconstructs existing systems than empirical research that generates guidelines for informing the design of such systems. Our work is situated in this latter realm and explores the question: How does an embodied view of cognition (and interaction) inform the way we design interaction for hybrid physical and digital environments?

As computing becomes embedded in the physical environment, one important aspect to understand is how to design interfaces that can support users to enact appropriate input actions and understand the relationships between these actions and digitally mediated output responses. How will such systems be made comprehensible to users? Svanæs (2001) suggests that the emergence of tangible user interfaces, context-aware devices and mobile computing requires new interface metaphors, and it is here that we turn our attention. The work presented in this paper explores the potential benefits and limitations of incorporating embodied metaphors in the interactional layer of an augmented audio environment. We focus on the interaction layer rather than the interface representation, to move away from a Cartesian model of cognition as information processing towards a phenomenological model of cognition as perception requiring action as described by Merleau-Ponty (2002).

In this paper we provide a description of The Sound Maker, an interactive audio environment that was designed to leverage embodied schemata in the interactional metaphor that relates input actions to output responses. The experiment was conceived to look for evidence of performance (tacit knowledge), explanatory (explicit knowledge) and experiential benefits of an embodied interaction model compared to a non-embodied model in the same interactive audio environment. We conclude the paper with suggestions for high level guidelines relevant to the design of interactive systems that rely on embodied, physical interaction.

2. Related work

This section describes the role that interaction models play in interactive audio environments. We follow this with a summary of the theoretical foundation required to understand the inclusion of an embodied metaphor in the interactional model. The section concludes with a discussion of previous uses of metaphors in graphical user interfaces and ubiquitous computing environments.

2.1. Interactive audio environments

In the 1970's and 1980's computer artists such as Myron Krueger and David Rokeby created computer-controlled light and sound environments that responded to the people within them. Recent advances in sensor systems and computer vision algorithms along with the maturation of computer generated sound systems have supported the development of a variety of interactive audio systems, which create sounds in response to body movement and physical interaction. Interactive audio systems have been developed as research instruments to explore specific phenomena including the social aspects of interaction; discovery-based play; naïve understandings of sensors; expressive intention; social participation in space; the promotion of physical activity in children (respectively, Aoki et al., 2002; Ferris and Bannon, 2002; Andersen, 2004; Fenza et al., 2005; Williams et al., 2005; Zigelbaum et al., 2006). There are no audio systems to date that have been developed to explore the role of metaphor in an interactional space.

Interactive audio environments differ by the type of sensor used and their musical capabilities (Morales-Manzanares et al., 2001). We focused on a system that used camera vision to track continuous full-body movement to control percussive audio output. Wrinkler provides a good overview of movement sensing for interactive music composition (1995) which informed our work. Movement data may be mapped to musical parameters including volume, tempo, rhythm, pitch and beat. Further, movement data may be selected, scaled or filtered before it is used as parameter input to compositional algorithms. Performers' individual or group actions can be translated into immediate, delayed or cumulative musical responses.

With an acoustic instrument the playing interface is often integrated with the sound source. For example, with a violin the strings are part of both the control and the sound production mechanisms. This is not so with electronic musical interfaces. The interface and control mechanism are usually completely separate from the sound production source. This means that the interaction model, that is the mapping layer between control (input actions) and sounds (output responses), must be explicitly defined. Hunt et al. state that by altering this mapping layer and keeping the interface itself and sound source constant, the entire nature of the environment (or instrument) is changed (2002). Our implementation supports the inclusion of different interaction models (mapping layers) in the same audio system to facilitate an experimental comparison as described in Section 3.1 below.

2.2. Metaphor theory

In language and thought metaphors help us understand one thing in terms of another. The word metaphor comes from the Greek word 'metaphora' meaning 'transfer'. Originally, metaphor was understood as a figurative expression, which interpreted a thing or action through an implied comparison with something else. Aristotle described metaphors in these comparative terms in the Rhetoric (Aristotle, 2004). This conception of metaphor held until the 1930s when it was opposed by the interactionalist definition that proposed that one thing is not simply a substitute or comparator for another but that metaphor is a process, which brings two terms into interaction with each other. In this way, the brain uses one concept as a filter or extractor for another (Black, 1962). In the late 1970s, John Searle rejected both the comparison and interaction theories of metaphor (Searl, 1979). He offered an understanding of metaphor based on the speaker's utterance meaning. He claimed that metaphorical utterances work not because a certain interaction of words produces a change in the meaning of both elements but because the speaker's meaning differs from their literal usage.

In the 1980s Lakoff and Johnston proposed a subclass of metaphor, a *conceptual metaphor* (1980). They suggested that conceptual metaphors run deeper than simple linguistic conventions. Rather than just an interaction of two words, a metaphor is the interaction between a target domain and a source domain that involves an interaction of schemas or concepts. As such, metaphors are systematic thought structures. Johnson argued that metaphor is one of our primary cognitive structures for ordering experience (Johnson, 1987). He claimed that metaphors arise unconsciously from experiential gestalts relating to the body's movements, orientation in space, and its interaction with objects. He called these fundamental gestalts *embodied schemata*. Conceptual metaphors extend embodied schemata to structure and organize abstract concepts. We call metaphors that are based on embodied schemata and operate at a preconscious or sensori-motor level of experience, *embodied metaphors*.

Embodied metaphors conceptually extend embodied schemata through the linking of a source domain that is an embodied schema and a target domain that is an abstract concept. For example, the body's general upright position in space creates a verticality schema that results in various spatial metaphors based on a vertical hierarchy (Lakoff and Johnson, 1980). Humans experience a physical world in which sticks added to a pile or water added to a container results in the level increasing. These interactions with the physical environment support the association *up* as *more* (as opposed to *down* as *more*). Embodied metaphors based on spatial experiences are called *orientational metaphors*. Orientational metaphors give an abstract concept a spatial orientation. For example, HAPPY IS UP and SAD IS DOWN. These metaphors lead to expressions in English such as "I'm feeling *up* today". Orientational metaphors are often used to interpret music (Budd, 2003). For example, "The music *lifted* me *up*." However, the use of orientational metaphors in understanding music is largely related to the emotional impact or content of the music, rather than individual concepts related to aspects of musical sounds (e.g., amplitude, tempo). We began our investigation with the assumption that *ontological* metaphors (described below) may be more appropriate than orientational metaphors for designing a movement-based system because ontological metaphors may be related to qualities of movement of objects (including humans) rather than relationships in space.

An *ontological* metaphor represents an abstract concept as something concrete and physical such as an object, person, body or substance in the environment (Lakoff and Johnson, 1980). Understanding our experiences in this way allows us to treat parts of our experiences as discrete entities, objects or substances of a uniform kind that can be referred to, categorized, grouped, quantified and qualified. Even when things are not discretely bounded, we refer to them in this way. For example, INFLATION IS AN ENTITY allows us to reason about the abstract concept of inflation as if it was a discrete entity. We might say, "Inflation is rising". Another example is MUSIC IS A SUBSTANCE. We might say "The music *flowed* into the auditorium". Alternatively, we can interpret music through the metaphor MUSIC IS PHYSCIAL BODY MOVEMENT (Barker, 1989; Jensenius, 2007). For example, "The music *raced* to *its* conclusion." It is this metaphor that we focus on as discussed in Section 3.1.4. Other musical concepts such as musical processes (e.g., melody, harmony, rhythm) and musical works themselves are also often understood through spatial and physical metaphor but these are not the focus of our study.

2.2.1. Children and metaphor

Direct physical interaction with the world is a key component of cognitive development in childhood. Piaget began a long tradition of thought that suggests that cognitive structuring through schemata accommodation and assimilation requires both physical and mental actions (Piaget 1952). Piaget proposed that children develop abstract conceptions based in part on the extension of concrete, physical schemata. Many embodied metaphors operate at the preconscious level of awareness (Lakoff and Johnson, 1980; Hurtienne and Israel, 2007). In terms of learning, psychologists have shown that children may learn something new and intuitively put it into action before they are able to consciously verbalize it (Myers, 2002). For example, Goldin-Meadow's studies of children and gesture demonstrated this effect (2005). Children often have preconscious knowledge of embodied schema and related metaphorical concepts, and are able to act on this knowledge, yet are not be able to verbalize it. As they age, preconscious knowledge often becomes conscious and explicit. In our empirical study we theorized that we would see a larger discrepancy between performance (tacit) and verbalized (explicit) knowledge in children than in adults (Antle et al., 2008).

2.3. Graphical user interface metaphors

The use of metaphor in graphical user interface (GUI) design has been well covered in past research (e.g., Carroll and Thomas, 1982; Laurel, 1986; Wozny, 1989; Erickson, 1990; Svanæs, 2001; Blackwell, 2006). In GUIs many elements of the interface are modeled on objects or actions taken from the physical world. Metaphor is commonly used as the basis for interface representations that are created to help the user understand the abstract workings of computer systems. The desktop metaphor, common in personal computing, is often given as a classic example (although Imaz and Benyon (2007) suggest that it is a conceptual blend rather than a metaphor). Interface elements are modeled on common desktop objects (e.g., file folder, wastebasket) and actions on those objects (e.g., opening files, throwing out files into a wastebasket). While most human computer interaction reference books include an entry for metaphor and give examples of the use of metaphor in user interfaces, the potential benefits of using metaphor are not uncontroversial. As Blackwell points out, the understandings of the benefits and limitations of designing metaphor-based user interfaces have changed over time (Blackwell, 2006).

2.4. Ubiquitous computing interface metaphors

The compatibility of mappings in physical interfaces has a long history beginning with the Dutch physiologist Franz Donder's recognition of compatibility effects in the 1860's. These ideas resurfaced with Fitts' work in the 1950's when he made a case for exact spatial duplication in control interfaces (Proctor and Vu, 2006). The concept of metaphorical mappings in tangible user interfaces (TUIs) was introduced by Svanæs in his doctoral thesis (2000) but his focus was primarily on GUIs. Koleva et al. (2003) discussed possible mappings in TUIs but did not introduce the concept of metaphor directly. Fishkin (2004) introduced a taxonomy for the classification of TUI research, which used metaphor and embodiment as its two dimensions. The work was descriptive and relied on an alternative interpretation of the term embodiment, that which refers to the degree to which the technology is embedded in an object. Blackwell et al. (2005) suggested that solid diagrams can be used to understand the relationships between physical and abstract problem elements in TUI prototyping. Klemmer et al. (2006) stressed the importance of compatibility of mappings where a spatial equivalent is possible and introduced the suggestion to use metaphor for non-spatially equivalent mappings. Hurtienne and Israel (2007) presented a taxonomy for classifying TUIs based on metaphorical extensions of image schemas. However, neither Klemmer et al. nor Hurtienne and Israel provided guidelines for the use of metaphor in mappings. Empirical work is still needed to understand the benefits and limitations of using metaphors in the design of ubiquitous computing applications.

3. Research design and methods

To investigate the potential benefits and limitations of utilizing an interactional model based on an embodied metaphor, we built an interactive audio environment called The Sound Maker. The "sound maker" is the user who controls the system through their movements. In order to facilitate a comparative experimental approach, The Sound Maker was implemented with two different interaction models. One model instantiated an embodied metaphor and the other did not. The independent variable was the interaction model. For reasons of ecological validity we used a collaborative, paired condition. Using a between subjects design to avoid order effects, 20 pairs of adults and 20 pairs of children completed a series of structured exercises followed by a composition exercise using only one model of the system.

3.1. System design: The Sound Maker

The Sound Maker is a room sized interactive audio environment. Users control percussive sounds (four instruments) and associated sound parameters (volume, tempo, pitch and rhythm) through continuous full-body movement in the space.

3.1.1. System inspiration

Inspiration for our interactive audio environment came from Dalcroze Eurhythmics (Jaques-Dalcroze, 1972). In the 1930's Jaques-Dalcroze proposed an alternative form of music education for children. In his approach, the primary form of knowing was movement-based rather than relying on an abstract and conceptual mode of teaching music theory, which was then the dominant mode of music education in Western cultures. Instead of focusing on teaching techniques necessary to play an instrument, Dalcroze Eurhythmics aimed to develop bodily knowing and awareness of the physicality of performing music. Bodily knowing relies on tacit knowledge, pre-reflective knowing. A competent performance occurs when the performer has internalized actions and has no further need to focus on body parts, as a beginning performer might. This reliance on and development of tacit knowledge in Eurhythmics is an appropriate domain in which to explore metaphors that support embodied interaction. For an excellent description of Dalcroze Eurhythmics see Juntunen and Hyvönen (2004).

3.1.2. Design criteria

Our goal was to create a system that we could use as an experimental test-bed to look for evidence that leveraging embodied knowledge in interaction design supports users to learn to control an interactive audio environment. There are various kinds of metaphors that are used to understand different aspects of music. For an interactive environment that relies on full-body movement we propose that the ontological metaphor MUSIC IS (BODY) MOVEMENT is appropriate as discussed in Section 2.2.

The major design goal was to create a system that related physical movement to changes in output sound parameters (e.g., amplitude, tempo, pitch). The primary criterion for the system was that the interface would be distributed in a space that facilitated movement (input) and produced variable sound responses (output). A second criterion for the system was that it had to be usable by both adults and children. Previous work by the authors provided evidence that children in this age range can perceive scaled differences in volume, tempo and pitch in a responsive environment (Droumeva et al., 2007). A third major criterion was that the system had to support the inclusion of different interactional models without changing the interface or the sound source.

Based on pilot studies and in order to maximize the potential to leverage preconscious knowledge rather than prior music learning or analytical ability, we had a constraint that the system should give no immediately perceivable cues to its usage. For example it should avoid a spatial layout that mimicked the layout of musical instrument (e.g., rectangular layout of a piano or keyboard where pitch varies with distance from centre).

3.1.3. The Sound Maker interactive audio environment

Our interactive audio environment, The Sound Maker, addresses the design goals by using a camera vision system to track pairs of users' movements in a rectilinear space divided into four unmarked quadrants. The system relates qualities of movement to discrete changes in percussive audio output. Users control the sequencing of percussive sounds and the change of musical parameters of those sounds through their collaborative body movements in the space. Videos of adult participants learning to use The Sound Maker can be found at (http://www.antle.iat.sfu.ca/Embodied-Metaphor).

The Sound Maker environment was implemented using a Panasonic WV-CP470 3CCD camera, which fed video data through a Video Data DAC100 video digitizer to a colour tracking system programmed in Max/MSP/Jitter on a Mac Laptop. The camera was set to provide a top view of an area 5.1 m $\times$ 4.5 m ($17' \times 15'$) that enabled the tracking of two participants within this footprint. The tracking system provided separate position data for each participant at a rate of 25 frames per second and transmitted this data to a second Mac laptop through a TCP/IP LAN connection using a FS105 NETGEAR Ethernet Switch. A Max/MSP patch ran on the second laptop, which analyzed the position data (as shown in Fig. 1). The system used sensed data to infer users' *speed* (i.e., rate of change of user position), the amount of *activity* in their movements (e.g., waving arms and stomping feet versus walking stiffly), the relative position or *proximity* of each user in the space (e.g., moving closer together), and the *flow* of their movements (e.g., synchronous/smooth versus asynchronous/choppy). Speed and activity can be distinguished by the following example. When a high level of *activity* occurs and the participant is standing in one place (e.g., running on the spot), the *speed* is zero since speed was defined as the rate of change of position (in any direction).

These movement characteristics were then mapped to control parameters: *volume* (loud, soft), *tempo* (fast, slow), *pitch* (high, low) and *rhythm* (rhythmic, chaotic) of four sequencers. Each sequencer generated a separate rhythmic pattern based on the data being mapped to the control parameters and used an individually assigned sound. Percussive sounds (marimba, celesta, pizzicato viola and woodblock) were chosen to ensure perceivability of a wide variety of tempos and rhythms. The output of each sequencer was then wired to one of four Yamaha MSP5A speakers.

3.1.4. Interaction models

The MUSIC IS MOVEMENT metaphor suggests that aspects of music may be treated as human body movement. Changes in sound parameters have physical origins. However, the concepts used to understand sound parameters are often quite abstract to people not trained in music or acoustics. The concrete human body movement source domain helps us to understand the more abstract musical sounds target domain.

To avoid our own biases, we conducted two rounds of semi-structured interviews with experts in movement and music: four dancers and/or choreographers. We first asked them to generate potential movement to sound parameter mappings. The first results showed quite a bit of variation in the ways they envisioned input movements. We created a set of input movements informed by Dalcroze Eurhythmics and based on the interview results and our observations from two pilot studies. We constrained inputs to qualities and quantities of movement (e.g., moving quickly, moving actively) rather than specific types of movements (e.g.,

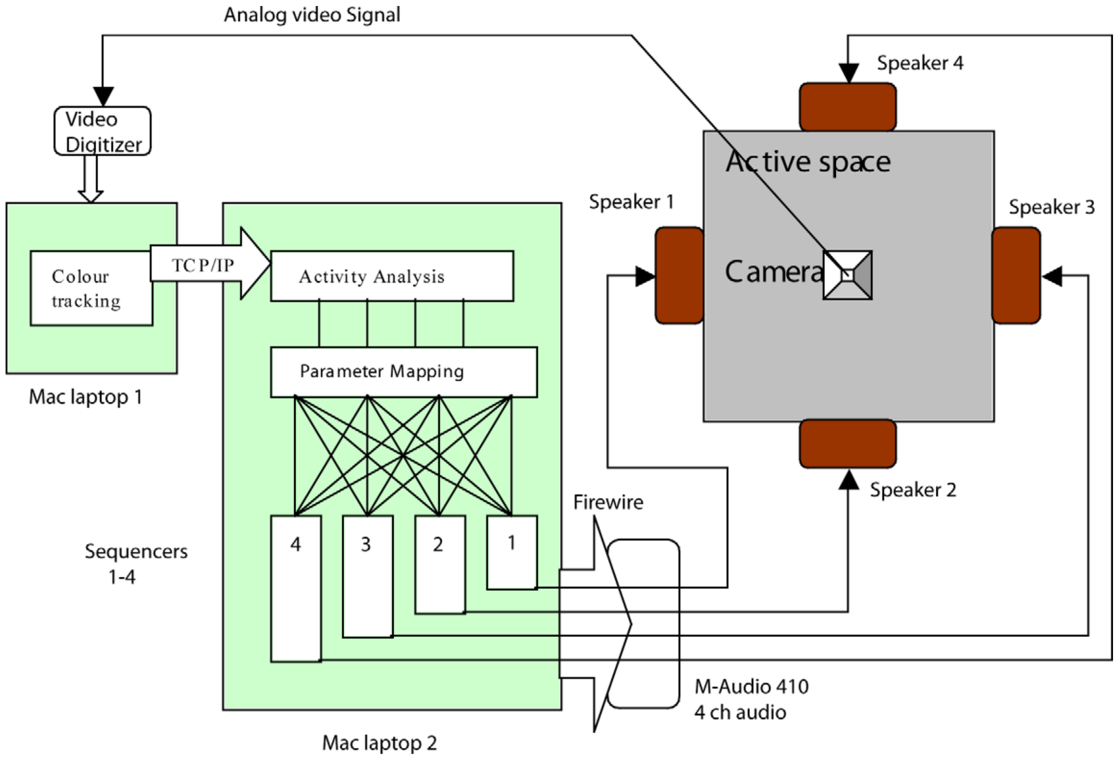

Fig. 1. The Sound Maker system infrastructure.

jumping, stomping) to avoid inconsistency in enactment. We eliminated movements that were difficult to sense (e.g., moving quietly). We then asked interviewees to match this set of qualities and quantities of body movement with a range of sound parameters. The results from the second interviews produced general agreement on viable mappings for volume, tempo, pitch and rhythm. For example, tempo was associated with speed of movement through (or around) a space. Pitch was associated with movement up and down in 3D space or towards and away from in 2D space. We eliminated parameters for which there was no expert agreement or children in this age range have difficulty perceiving (e.g., timbre) (Droumeva et al., 2007). The movement experts also validated the polarity of the mappings. Polarity refers to the direction of gradient of change. The final mappings for the embodied metaphor-based interaction model are shown in Table 1. The final mappings for the non-metaphorical model were chosen in opposition to expert opinions and are shown in Table 2.

Table 1
Embodied metaphorical mappings.

Movement	Parameter	Mappings
Speed	Tempo	Fast is fast; slow is slow
Activity	Volume	More is loud; less is quiet
Proximity	Pitch	Near is high; Far is low
Flow*	Rhythm	Smooth is rhythmic; choppy is chaotic

* Only adults did tasks based on these mappings (due to session time constraints with children).

Table 2
Non-metaphorical mappings.

Movement	Parameter	Mappings
Flow	Tempo	Smooth is fast; choppy is slow
Proximity	Volume	Far is quiet; near is loud
Speed	Pitch	Slow is high; fast is low
Activity*	Rhythm	High is rhythmic; low is chaotic

* Only adults did tasks based on these mappings (due to session time constraints with children).

3.2. Hypotheses

Our research design was driven by the specific question: What are the potential benefits (and limitations) of incorporating embodied metaphors in the interactional layer of an augmented audio environment?

Based on the theoretical work previously discussed we would expect that both adults and children would find it more efficient, effective, and enjoyable to use the metaphor-based version of The Sound Maker than the non-metaphor-based version. We would also expect that these benefits might be more noticeable when children perform using the system compared to when they are asked to verbally explain how the system works. Our six hypotheses follow from these expectations. However, the complexities seen in work with GUI metaphors lead us to also expect the unexpected. In addition, we recognize the limitations of our study, which focuses on a short and initial snapshot of novice users' experiences with The Sound Maker.

H1: All participants (adults and children) will require less practice time to learn to use the metaphor-based version than the non-metaphor-based version of The Sound Maker.
H2: All participants will perform (demonstrate) specific sound sequences (patterns) more accurately using the metaphor-based version than the non-metaphor-based version of The Sound Maker.
H3: All participants will verbally explain how to create specific sound sequences more accurately using the metaphor-based version than the non-metaphor-based version.
H4: Adults will be able to perform and explain equally well how to create specific sound sequences with the metaphor-based version.

H5: Children will be able to better perform than verbally explain how to create specific sound sequences with the metaphor-based version.

H6: All participants will find the metaphor-based version more enjoyable to use than the non-metaphor-based version.

3.3. Participants

The study was comprised of sessions with 20 pairs of adult volunteers (total 40 participants) of both genders (16 males, 24 females), aged 18–40 years old, recruited from the urban Simon Fraser University campus. We also conducted sessions with 20 pairs of child volunteers (total 40 participants) of both genders (20 males, 20 females), aged 7–10 years old, recruited from an urban science centre. No previous musical experience was required. Participants with significant musical or acoustic training were not included in the study. All participants used computers daily or weekly. There were no significant differences in adult's or children's preference ratings for music or physical activity between groups. Participants were randomly grouped in gender-matched pairs (where possible).

3.4. Procedure

Recognizing, mimicking and creating simple sound sequences with variations in volume, tempo, pitch and rhythm are common activities used to teach young children music (Juntunen and Hyvönen, 2004). Since participants were not required to have any musical training, beginner level exercises were chosen for both adults and children. Paired participants were asked to work together to create sound sequences by moving their bodies in the space. This type of movement-based exercise is common in the Dalcroze Eurhythmics approach to music education (Jaques-Dalcroze, 1972). Children were given a subset of the adult tasks as described below.

After a free play session, the participants were given a series of four tasks in which they were asked to create specific sound sequences by varying a *single* parameter (volume, tempo, pitch, rhythm). For example, in the "volume" task, they were asked to make a sound sequence where the volume varied from loud to quiet and back to loud. Results from pilot studies with both adults and children helped us calibrate sound output scales and ensured that sensed movements created changes in sounds that children could perceive. The fifth and sixth tasks involved creating a sound sequence by varying two parameters at once (volume and tempo, pitch and rhythm). In order to keep the total session time under an hour, the children only completed the first three and fifth tasks and the adult verbal explanation data was collected for a subset of the tasks. After all structured tasks were completed, all participants were given the opportunity to compose their own sound sequence, which they then demonstrated and explained to the facilitator.

3.5. Measures

In order to provide evidence for claims of benefit we collected and analyzed several forms of data, both quantitative and qualitative. For each task, we recorded the time it took the participants to practice creating sound sequences with a maximum of 10 minutes per task. Practice time may be used as a measure of the difficulty in learning a system. Two researchers rated the accuracy of participant's performed sequences and their verbal explanations for each task (correct, partially correct or incorrect). Consensus was achieved for all ratings based on the set of rules for coding accuracy established from the pilot studies. An example of a partially correct solution for performance of the tempo task (fast–slow–fast) was if a pair moved quickly and increased the tempo reliably but did not create slow tempo reliably or did not alternate fast and slow tempo variation. For statistical purposes, performance accuracy codes were scored two for correct, one for partially correct, and zero for incorrect. Participants individually completed a post session questionnaire based on the Intrinsic Motivation Inventory subscales for Enjoyment and Interest and Perceived Competence based on the author's previous work (Ryan, 2006; Xie et al., 2008). They also gave ratings for individual statements related to ease of learning, intuitiveness of learning and amount of concentration required to learn to use the system.

Two researchers took notes throughout the sessions. For example, we noted any verbal explanations participants made about their sound sequences while they were demonstrating, and noted any discrepancies between how they performed and what they verbalized. We recorded participants' responses when they were asked what they liked and disliked about their experiences after the session. We also video taped all sessions for later close video analysis and validation of accuracy ratings.

4. Results

4.1. Quantitative results

Descriptive and inferential statistics were used to analyze time, accuracy data, and questionnaire ratings. The time data was non-normally distributed and had unequal variances. The accuracy data was coded to ordinal values of two to zero and did not conform to a normal distribution. Questionnaire data showed a skewed distribution typical of Likert data. As a result, the non-parametric Mann–Whitney U test, which is suitable for ordinal or continuous data, was used for all quantitative variables.

4.1.1. Hypothesis one: adult's and children's practice time

Descriptive statistics of the practice times for adults and children of each condition are shown in Table 3. Although there is a trend for all participants to take longer with the non-metaphor-based system on all tasks, Mann–Whitney U tests revealed that participant's practice times were not significantly different between the two groups, with the exception of the adult pitch task. Adult's pitch task practice time was significantly shorter for the metaphor-based model ($U = 16.0$, $p < .01$). Except for this anomaly (discussed in Section 4.2.1), there is no inferential evidence to support hypothesis one as operationalized.

It is evident that task times do not reflect task difficulty since the results for accuracy (below) suggest a strong pattern of poor accuracy for the non-metaphor group and Pearson's correlation between time and accuracy data showed no consistent pattern of correlations. The lack of significance may be largely due to the large variation in practice times within both conditions as seen in the standard deviations. From our observations we suggest that personality factors may play an important role in determining how long participants are willing to spend learning the system and practicing each task.

Histograms of the children's practice times for the metaphor-based group show a pattern in which most pairs complete the volume, tempo and volume-tempo task relatively quickly but several pairs take much longer. This could be interpreted in light of our observation about personality factors. That is, several pairs of children practiced creating specific sound sequences repeatedly, perhaps working to perfect them. Histograms for adult practice times for the metaphor-based group show this same pattern only for the more difficult tasks (e.g., rhythm and mixed tasks). Histograms for the children's and adult's practice times for the non-metaphor-based group do not show this pattern.

Table 3
Descriptive statistics for practice time data (seconds) a. Adults; b. Children.

a. Adults	Vers*	N	Mean (s)	Std Dev	b. Children	Vers*	N	Mean (s)	Std Dev
Task 1: Volume	Metaph	10	123	108	Task 1: Volume	Metaph	10	235	239
	Non	10	71	59		Non	10	253	181
Task 2: Tempo	Metaph	10	269	193	Task 2: Tempo	Metaph	10	107	158
	Non	10	253	199		Non	10	195	142
Task 3: Pitch	Metaph	10	106	67	Task 3: Pitch	Metaph	10	112	86
	Non	10	267	143		Non	10	123	79
Task 4: Rythm	Metaph	10	322	146	Children did not do this task				
	Non	10	392	167					
Task 5: VolTempo	Metaph	10	194	132	Task 5: VolTempo	Metaph	10	88	129
	Non	10	277	175		Non	10	179	147
Task 6: PitchRhy	Metaph	10	233	162	Children did not do this task				
	Non	10	371	196					

* Metaph = mapping layer utilizes embodied metaphor; Non = mapping layer does not utilize embodied metaphor.

It is also possible that longer mean task times indicate engagement rather than task difficulty. A larger sample size might reduce the impact of individual differences.

4.1.2. Hypothesis two: adults' and children's performance accuracy

The frequency counts of adult performance accuracy codes yielded expected results as shown in Tables 4a–10a. Adults performed more accurately on all tasks, except the volume task, when using the metaphor-based system. For the volume task both groups performed equally well (Table 4a). The percentage of correct performances across all tasks for the metaphor-based group is 80% compared to 15% for the non-metaphor-based group. If we remove the volume task that was anomalous (discussed in Section 4.2.1) then we find that adults performed correctly in 80% of cases with the embodied metaphor version compared to only 2% with the non-embodied version. Analysis using Mann–Whitney tests showed significant differences between the two adult groups in the performance accuracy scores across all tasks except the volume task. Specifically, Mann–Whitney tests for the tempo task showed a significant mean difference between groups ($U = 5.5$, $p < .0001$); pitch task ($U = 10.0$, $p < .01$); rhythm task ($U = 1.0$, $p < .0001$); volume-tempo task ($U = 0.0$, $p < .0001$) and pitch-rhythm task ($U = 6.5$, $p < .001$).

The frequency counts of the children's performance accuracy codes yielded expected results on all tasks except the pitch task as shown in Tables 4b–10b. The percentage of correct performances across all tasks for the metaphor-based system is 63% compared to 8% for the non-metaphor-based system. If we remove the pitch task that was anomalous (discussed in Section 4.2.1) then we find that children performed correctly in 80% of cases with the embodied metaphor version compared to only 7% with the non-embodied version. Analysis using Mann–Whitney tests showed significant differences between the two children's groups across all tasks except the pitch task. Specifically, Mann–Whitney tests for the volume task showed a significant mean difference between groups ($U = 25.0$, $p < .05$); tempo task ($U = 8.0$, $p < .001$); and volume-tempo task ($U = 2.0$, $p < .0001$). There was no significant difference between groups for the pitch task. Participants in both groups performed poorly on this task.

Table 4a
Volume task: counts of accuracy codes.

Adults	Vers	Correct	Partial	Incorrect
Volume perform	Metaph	8	2	0
	Non	8	2	0
Volume verbal	Metaph	1	8	1
	Non	8	1	1

Table 4b
Volume task: counts of accuracy codes.

Children	Vers	Correct	Partial	Incorrect
Volume perform	Metaph	8	0	2
	Non	2	5	4
Volume verbal	Metaph	3	3	4
	Non	2	1	7

Table 5a
Tempo task: counts of accuracy codes.

Adults	Vers	Correct	Partial	Incorrect
Tempo perform	Metaph	7	3	0
	Non	0	1	9
Tempo verbal	Metaph	7	3	0
	Non	0	1	9

Table 5b
Tempo task: counts of accuracy codes.

Children	Vers	Correct	Partial	Incorrect
Tempo perform	Metaph	8	1	1
	Non	0	3	7
Tempo verbal	Metaph	8	1	1
	Non	0	0	10

Table 6a
Pitch task: counts of accuracy codes.

Adults	Vers	Correct	Partial	Incorrect
Pitch perform	Metaph	9	1	0
	Non	0	0	10
Pitch verbal	Metaph	8	1	1
	Non	0	0	10

Table 6b
Pitch task: counts of accuracy codes.

Children	Vers	Correct	Partial	Incorrect
Pitch perform	Metaph	1	1	8
	Non	1	0	9
Pitch verbal	Metaph	2	0	8
	Non	0	1	9

Overall, we have strong evidence that participants were able to more accurately perform specific sound sequences using the metaphor-based interaction model than the non-metaphor-based model.

Table 7
Rhythm task: counts of accuracy codes.

Adults	Vers	Correct	Partial	Incorrect
Rhythm Perform	Metaph	7	3	0
	Non	0	1	9
Rhythm Verbal	Metaph	5	4	1
	Non	0	1	9
Children did not do this task				

Table 8a
Volume-Tempo task: counts of accuracy codes.

Adults	Vers	Correct	Partial	Incorrect
VolTemp Perform	Metaph	10	0	0
	Non	0	3	7
VolTemp Verbal	Metaph	NA	NA	NA
	Non	NA	NA	NA

Table 8b
Volume-Tempo task: counts of accuracy codes.

Children	Vers	Correct	Partial	Incorrect
VolTempo perform	Metaph	8	2	0
	Non	0	2	8
VolTempo verbal	Metaph	3	7	0
	Non	0	0	10

Table 9
Pitch-Rhythm task: counts of accuracy codes.

	Vers	Correct	Partial	Incorrect
VolTemp Perform	Metaph	7	3	0
	Non	1	0	9
VolTemp Verbal	Metaph	NA	NA	NA
	Non	NA	NA	NA
Children did not do this task				

Table 10a
All tasks: percentages for accuracy codes.

Adults	Vers	Correct (%)	Partial (%)	In-correct (%)
All tasks perform	Metaph	80	20	0
	Non	15	12	73
All tasks verbal	Metaph	53	40	7
	Non	20	8	72

Table 10b
All tasks: percentages for accuracy codes.

Children	Vers	Correct (%)	Partial (%)	Incorrect (%)
All tasks perform	Metaph	63	10	27
	Non	8	25	67
All tasks verbal	Metaph	40	28	32
	Non	5	5	90

4.1.3. Hypothesis three: adults' and children's verbal explanation accuracy

The frequency counts of the adult verbal explanation accuracy codes yielded expected results, as illustrated in Tables 4a–10a. Adults explained how to control the system more accurately on all tasks, except the volume task, using the metaphor-based system. For the volume task, the frequency counts for correct explanations are higher for the non-metaphor group (Table 4a). Observational notes indicated that participants often thought that volume was mapped to speed rather than activity (i.e., amount of movement). Since movements, which involve fast traversal of space (speed), tend to involve a lot of movement (activity), it is easy to see how the two might be confused.

The percentage of correct verbal explanations across all tasks for the metaphor-based system is 53% compared to 20% for the non-metaphor-based system. If we again remove the volume task then we find that adults gave correct verbal explanations in 67% of cases with the embodied metaphor version compared to 0% with the non-embodied version. Analysis using Mann–Whitney tests showed expected significant differences between the two adult groups in verbal explanation accuracy scores across all tasks except the volume task. Specifically, Mann–Whitney tests for the tempo task showed a significant mean difference between groups ($U = 1.0$, $p < .0001$); pitch task ($U = 5.0$, $p < .0001$); and rhythm task ($U = 7.5$, $p < .0001$).

The frequency counts of the children's verbal explanation accuracy codes yielded expected differences on all tasks except the pitch task, as shown in Tables 4b–10b. The percentage of correct verbal explanations across all tasks for the metaphor-based system is 40% compared to 5% for the non-metaphor-based system. Analysis using Mann–Whitney tests showed significant differences between the two children's groups in verbal explanation accuracy scores for the tempo and volume-tempo tasks. Specifically, Mann–Whitney tests for the tempo task showed a significant mean difference between groups ($U = 5.0$, $p < .0001$) and volume-tempo mixed task ($U = 0.0$, $p < .0001$). There was no significant difference between groups for the volume or pitch tasks.

Overall, we have some evidence that participants were more accurately able to verbally explain how to create specific sound sequences using the metaphor-based interaction model. However, there is a dependency on task. From observational notes, we suggest that for adults, changes in sound volume provided a salient perceptual clue to system functionality regardless of interaction model. We discuss this under the qualitative theme *Discoverability* below in Section 4.2.1. In addition, adults often enacted movements towards or apart from each other, a behavior rarely seen in the children's pairs. This may explain the lack of difference between children's groups on the pitch task since pitch was mapped to proximity in the metaphor-based system.

4.1.4. Hypothesis four: adults' performance versus verbal explanation accuracy

A comparison of performance and verbal accuracy on all tasks revealed that in the embodied metaphor-based system adults *correctly* performed tasks in 80% of the cases and verbally explained their sound sequences in only 53% of the cases. There was a small significant difference ($U = 615.5$, $p < .05$). However, if we remove the volume task (since speed and activity were often conflated), there is no significant difference.

4.1.5. Hypothesis five: children's performance versus verbal explanation accuracy

A comparison of performance and verbal accuracy on all tasks revealed that in the embodied metaphor-based system children *correctly* performed tasks in 63% of the cases and verbally explained their sequences in only 40% of the cases. The difference is not significant. In a comparison excluding the pitch task (since children could neither perform nor explain it), children *correctly* performed the remaining tasks in 80% of the cases and verbally explained their sequences in only 47% of the cases. A Mann–Whitney test comparing correct codes for tasks excluding pitch showed a significant difference ($U = 73.5$, $p < .01$). These results provide evidence that children may be able to perform sound sequences better than they can verbally explain when using the embodied metaphor-based system.

4.1.6. Hypothesis six: adult's and children's experience ratings

There were no significant differences in mean ratings for ease of learning, enjoyment or perceived competence scales between groups. While this result may seem surprising, one of authors has found a lack of experiential effects in a similar style of study (Xie et al., 2008). We suggest that participating in a scientific experiment of this nature is in and of itself, enjoyable regardless of outcome. In addition, participants were not told if they achieved accurate sequences for any tasks.

4.2. Qualitative results

Qualitative analysis of detailed observational notes and video revealed three salient themes, which we used to contextualize the quantitative findings and characterize how the participants interacted with the two different interaction models of The Sound Maker.

4.2.1. Discoverability

We categorized our interaction models as metaphor-based and non-metaphor-based depending on the nature of the mappings between input actions and system responses. While this characterization is important, it alone does not predict participant responses. A second dimension was induced from observation. We proposed that *discoverability* was a second dimension, which could be used to characterize the mappings in our interaction models and it significantly influenced participant responses. Discoverability means how likely it is that a particular group of participants will discover a mapping by chance.

The mapping of proximity to pitch in the metaphor-based system proved to be highly discoverable for adults. Adults often enacted a movement sequence where they started apart and then came close together, often to talk, and when they did so, they were rewarded with audible high pitch sounds. This may account for the significantly lower practice times in this group. The mapping of proximity to volume in the non-metaphor system also proved to be highly discoverable for adults as indicated by the accuracy scores. When the adult participants were far apart in space the system volume decreased quickly to silence. As soon as they began to move closer together, the volume increased. The strong visual cue of participant proximity combined with a salient auditory cue of silence made this mapping easily discoverable.

The inverse of discoverable was obscure. A mapping is obscure if it is unlikely that it will be revealed by chance actions. The mapping of proximity to pitch in the metaphor-based system proved to be obscure for child participants since they did not tend to move together or apart. In the non-metaphor-based system, the first two tasks did not reward high levels of movement speed or activity. When children began the third task (pitch) they had often settled into a pattern of slow movement since high speed and high activity levels had not been previously rewarded. This made it difficult for them to discover the mapping of speed to pitch. More study is needed but we stress the importance of discoverability of mappings in interaction design and note that what is discoverable depends on the user group and task order.

4.2.2. Structural isomorphism and perceivable feedback

Svænes (2001) suggests that identifying embodied metaphorical mappings between a concrete source domain and more abstract target domain that represent stable structural isomorphic relations between the two domains is beneficial. We found that participants had particular success with mappings that preserved structural isomorphisms between everyday movements and sound changes. For example, accuracy scores were high for tasks involving the mapping of movement speed to tempo.

We also observed that many participants enacted the kinds of physical movement qualities we envisioned. Participants commonly raced around the space, moved slowly in one place, and moved together in a synchronized way to elicit sound changes. In the metaphor-based system participant's initial movements were immediately rewarded with the desired changes in sound output. In the non-embodied metaphor version these first movements were not rewarded. Participants in the latter group often expressed some frustration or surprise and eventually resorted to other kinds of movements and actions. In some cases, they insisted that the movements they expected to work did work, even though sound feedback was clearly contrary to their claims.

We stress that structural isomorphism coupled with immediate perceptual confirmatory feedback are required to optimally leverage tacit embodied knowledge. We suggest that the benefits of using an embodied metaphor in the interactional model may be limited to guiding and constraining initial input actions if perceivable confirmatory feedback is not readily provided.

4.2.3. Duplicity of mappings

We observed participants interact with both systems in shared and unexpected ways. Our initial interviews with dancers and choreographers resulted in general agreement on mappings, however some duplicity of possible mappings was revealed in the process. For example, during the pilot studies we questioned if an increase in volume should be mapped to an increase in speed or to an increased amount of activity. The discrepancy between adult accurate performances and inaccurate verbal explanations for the volume task on the metaphor-based system gave direct evidence that there was not always a privileged mapping. Duplicity of mappings may need to be supported (e.g., mapping volume to both speed and activity since doing one increases the likelihood of doing the other). Support for users defining their own mappings may be also important to consider.

5. Design implications

We present the following suggestions for researchers and designers of ubiquitous computing systems that rely on full-body movement:

- Involving dancers and choreographers in design provides access to tacit body knowledge related to full-body movement.
- Understanding embodied schemata as source domain for metaphorical extensions to more abstract domains can be used as a resource to inform design.
- Utilizing interactional mappings based on embodied schema extended through conceptual metaphor provides some advantages to learning and usability aspects of an embodied style interface.
- Utilizing interactional mappings that preserve structural isomorphisms between lived experience and the target domain are particularly beneficial.
- Embodied interactional mappings must be easily discoverable by the intended user group to provide benefit.
- The advantage of embodied interactional mappings may not necessarily benefit users' explicit interpretation of the output representations (e.g., conceptualization of the sound parameters).
- Immediate perceivable feedback is required when users have performative knowledge but no explicit knowledge of appropriate input actions.
- The advantage of embodied interactional mappings may not extend past informing initial actions if perceivable confirmatory feedback is not provided.

- Supporting more than one mapping when duplicity exists in potential mappings will increase user's explicit understanding of the interaction model.

6. Conclusions

How does an embodied view of cognition inform the way we design interaction for hybrid physical and digital environments? How will users comprehend such systems? What are the potential benefits (and limitations) of incorporating embodied metaphors in the interactional layer of an augmented audio environment?

The Sound Maker research is an initial exploration of people's interactions with and through a full-body movement-based augmented environment. Our intention was not to create an innovative augmented audio environment but to build a system as a research instrument suitable to explore these questions in order to inform future research and design in ubiquitous computing. Our comparison of adult's and children's performance and explanations in learning to use The Sound Maker allowed us to determine and reflect on fundamental differences between interacting with a system that utilized an embodied metaphor in the interaction model and one that did not. We found that users were more effectively able to learn to use the system version with an interaction model based on an embodied metaphor. However, designing such interactional models requires consideration of interacting factors including the likely actions of the intended audience, the discoverability of mappings, which may or may not have duplicate interpretations, and the perceivability of supporting feedback. This study contributes empirically grounded understandings that can inform researchers and designers working within the embodied interaction paradigm of ubiquitous computing.

Acknowledgements

The authors wish to thank the anonymous reviewers for their thoughtful and supportive feedback. This research was funded by the National Science and Engineering Research Council of Canada (NSERC). Special thanks to Dr. Sandy Eix, the staff at the Telus World of Science, Nis Bojin, and all the SFU students and children who participated in the study.

References

Andersen, K., 2004. 'ensemble': playing with sensors and sound. In: Proceedings of Conference on Human Factors in Computing Systems, Vienna, Austria. ACM Press, pp. 1239–1242.

Antle, A.N., Droumeva, M., Corness, G., 2008. Playing with The Sound Maker: Do embodied metaphors help children learn? In: Proceedings of Conference on Interaction Design for Children, Chicago, IL, USA. ACM Press, pp. 178–185.

Aoki, P., Grinter, R., Hurst, A., Szymanski, M., Thorton, T., Woodruff, A., Voce, Sotto 2002. Exploring the interplay or conversation and mobile audio spaces. In: Proceedings of Conference on Human Factors in Computing Systems, Minneapolis, MI, USA. ACM Press, pp. 431–438.

Aristotle, 2004. Rhetoric. Kessinger Publishing.

Barker, A., 1989. Greek Musical Writings: Volume II, Harmonic and Acoustic Theory. Cambridge University Press, Cambridge, UK.

Black, M., 1962. Models and Metaphors. Cornell University Press, Uthaca, NY, USA.

Blackwell, A., 2006. The reification of metaphor as a design tool. ACM Transactions on Computer Human Interaction (TOCHI) 13 (4), 490–530.

Blackwell, A., Edge, D., Dubuc, L., Rode, J., 2005. Using solid diagrams for tangible interface prototyping. IEEE Pervasive Computing 4 (4), 74–77.

Budd, M., 2003. Musical movement and aesthetic metaphors. British Journal of Aesthetics 43 (3), 209–223.

Carroll, J., Thomas, J.C., 1982. Metaphors and the cognitive representation of computing systems. IEEE Transactions on Systems, Man and Cybernetics 12 (2), 107–116.

Clark, A., 1997. Being There: Putting Brain, Body and World Together Again. Bradford Books, MIT Press, Cambridge, MA, USA.

Clark, A., 1999. An embodied cognitive science? Trends in Cognitive Sciences 3 (9), 345–351.

Dourish, P., 2001. Where the Action Is. MIT Press, Cambridge, MA, USA.

Droumeva, M., Antle, A.N., Wakkary, R., 2007. Exploring ambient sound techniques in the design of responsive environments for children. In: Proceedings of Conference on Tangible and Embedded Interaction, Baton Rouge, LA, USA. ACM Press, pp. 171–178.

Erickson, T., 1990. Working with interface metaphors. In: Brenda, L. (Ed.), The Art of Human–Computer-Interface Design. Addison-Wesley, New York, NY, USA.

D. Fenza, M. Luca, S. Canazza, A. Rodia, 2005. Physical movement and musical gestures: a multilevel mapping strategy. In: Proceedings of Sound and Music Computing. Available from: <http://smc.afim-asso.org/smc05>.

Ferris, K., Bannon, L., 2002. . . .a load of ould boxology! In: Proceedings of Conference on Designing Interactive Systems, London, England. ACM Press, pp. 41–49.

Fishkin, K., 2004. A taxonomy for and analysis of tangible interfaces. Personal and Ubiquitous Computing 8 (5), 347–358.

Gallagher, S., 2005. How the Body Shapes the Mind. Clarendon Press, Oxford, UK.

Gibson, J.J., 1979. The Ecological Approach to Visual Perception. Lawrence Erlbaum, Houghton Mifflin, Boston, MA, USA.

Goldin-Meadow, S., 2005. Hearing Gesture: How Our Hands Help Us Think. First Harvard University Press, Cambridge, MA, USA.

Hunt, A., Wanderley, M., Paradis, M., 2002. The importance of parameter mapping in electronic instrument design. In: Proceedings of Conference on New Instruments for Musical Expression, Dublin, Ireland. National University of Singapore Press, Singapore, pp. 1–6.

Hurtienne, J., Israel, J., 2007. Image schemas and their metaphorical extensions: Intuitive patterns for tangible interaction. In: Proceedings on Tangible and Embedded Interaction, Baton Rouge, LA, USA. ACM Press, pp. 127–134.

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

Jaques-Dalcroze, E., 1972. Eurhythmics Art and Education. Benjamin Blom Inc., New York, NY, USA.

Jensenius, A.R., 2007. Ph.D. Thesis, Action–sound: Developing methods and tools to study music-related body movement. Available from: <http://www.arj.no/research/>.

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

Juntunen, M., Hyvönen, L., 2004. Embodiment in musical knowing: How body movement facilitates learning within Dalcroze Eurhythmics. British Journal of Music Education 21 (2), 199–214.

Klemmer, S., Hartmann, B., Takayama, L., 2006. How bodies matter: Five themes for interaction design. In: Proceedings of Conference on Designing Interactive Systems, University Park, PA, USA. ACM Press, pp. 140–149.

Koleva, B., Benford, S., Ng, K., Rodden, T., 2003. A framework for tangible user interfaces. In: Proceedings on Mobile HCI Conference: Physical Interaction (PI03) Workshop on Real World User Interfaces, Udine, Italy. ACM Press, pp. 1–5.

Lakoff, G., Johnson, M., 1980. Metaphors We Live By. Chicago Press, Chicago, IL, USA.

Laurel, B., 1986. Interface as mimesis. In: Norman, D.A., Draper, S.W. (Eds.), User Centered Systems Design. Lawrence-Erlbaum, Hillsdale, NJ, USA.

Merleau-Ponty, M., 2002. Phenomenology of Perception (Colin Smith, Trans.). Routledge, London, UK.

Morales-Manzanares, R., Morales, E., Dannenberg, R.F., Berger, J.F., 2001. SICIB: an interactive music composition system using body movements. Computer Music Journal 25 (2), 25–36.

Myers, D.G., 2002. Intuition: Its Powers and Perils. Yale University Press, USA.

Piaget, J., 1952. The Origins of Intelligence in Children. University Press, New York, NY, USA.

Proctor, R., Vu, K.-P., 2006. Stimulus Response Compatibility Principles: Data, Theory and Application. CRC Press, Taylor and Francis Group, Boca Raton, FL, USA.

Ryan, R.M., 2006. The Intrinsic Motivation Inventory (IMI). Available from: <http://www.psych.rochester.edu/SDT/measures/intrins.html>.

Searl, J., 1979. Expression and Meaning: Studies in the Theory of Speech Acts. Cambridge University Press, Cambridge, UK.

Svanæs, D., 2000. Ph.D. Thesis, Understanding interactivity: steps to phenomenology of human–computer interaction. Department of Computer Science, Trondheim, Norway. Available from: <http://dag.idi.ntnu.no/interactivity.pdf>.

Svanæs, D., 2001. Context-aware technology: a phenomenological perspective. Human Computer Interaction 16, 379–400.

Varela, F.J., Thompson, E., Rosch, E., 1991. The Embodied Mind. MIT Press, Cambridge, MA, USA.

Williams, A., Kabisch, R., Dourish, P., 2005. From interaction to participation: configuring space through embodied interaction. In: Proceedings of UbiComp, Toyko, Japan. ACM Press, pp. 287–304.

Wilson, M., 2002. Six views of embodied cognition. Psychonomic Bulletin & Review 9 (4), 625–636.

Wittgenstein, L., 1953. Philosophical Investigations. Blackwell, Oxford, UK.

Wozny, L.A., 1989. The application of metaphor, analogy and conceptual models in computer systems. Interacting with Computers 1 (3), 273–283.

Wrinkler, T., 1995. Making motion musical: gesture mapping strategies for interactive computer music. In: Proceedings on Computer Music Conference. Computer Music Association, San Francisco, CA, USA.

Xie, L., Antle, A.N., Motamedi, N., 2008. Are tangibles more fun? Comparing children's enjoyment and engagement using physical, graphical and tangible user interfaces. In: Proceedings of Conference on Tangible and Embedded Interaction, Bonn, Germany. ACM Press, pp. 191–198.

Zigelbaum, J., Millner, A., Desai, B., Ishii, H., 2006. BodyBeats: whole-body, musical interfaces for children. In: Extended Abstracts of Conference on Human Factors in Computing Systems. ACM Press, pp. 1595–1600.