

Human–computer-intuition? Exploring the cognitive basis for intuition in embodied interaction

Alissa N. Antle*, Greg Corness and
Milena Droumeva

School of Interactive Arts and Technology,
Simon Fraser University,
Surrey, British Columbia, Canada
E-mail: aantle@sfu.ca
E-mail: gcorness@sfu.ca
E-mail: mvdroume@sfu.ca
*Corresponding author

Abstract: One of the claimed benefits of embodied interaction is that it is an intuitive form of human–computer interaction. While this claim seems to be widely accepted, few studies explore the underlying cognitive mechanisms of intuition in the context of tangible and embedded interaction design. What is intuitive interaction? What makes an interface intuitive to use? We explore these questions in the context of a responsive auditory environment. We propose that intuitive interaction can be facilitated by instantiating an embodied metaphor in the mapping layer between movement-based input actions and auditory system responses. We search for evidence of benefit through a comparative study of the same responsive auditory environment implemented with and without an embodied metaphor in the interactional mapping layer. Qualitative findings about the complexities and limitations of designing intuitive interaction are summarised and the implications for the design of embodied interaction discussed.

Keywords: auditory environments; embedded interaction; embodied interaction; embodied schema; embodiment; image schema; interactive environments; intuitive interaction; metaphor; responsive environments; tangible user interfaces; whole body interaction.

Biographical notes: Alissa N. Antle is an Assistant Professor in the School of Interactive Arts and Technology at the Simon Fraser University. Her research focuses on embodied human–computer interaction and the design of tangible user interfaces and responsive environments. Her interactive work has been recognised by organisations including the New Media Invision Awards (GOLD), Parent’s Choice Foundation and SIIA Codie Awards. She earned her Bachelor’s degrees in Liberal Arts and Systems Design Engineering from the University of Waterloo and a PhD from the University of British Columbia.

Greg Corness is a Researcher and Artist Working with Embodied Interaction in Media Environments. He holds a Masters of Music in Computer Composition from the University of Victoria, Canada and is currently working on his PhD at the School of interactive Arts and Technology, Simon Fraser University,

Surrey, Canada. He Lectures on embodied interaction, interactive media and computer music. His research interests include media agents, generative systems and embedded computing. He has worked on projects for galleries, museums and live performance in Canada and the USA.

Milena Droumeva is a Doctoral Student in the area of Educational Technology at Simon Fraser University, studying the role of sound in learning environments, and particularly the notion of acoustic epistemologies. She completed her BA in Communications and received a Masters in the School for Interactive Arts and Technologies, SFU, where she explored a design framework for sound design in ubiquitous computing environments. She has also designed sound for a variety of systems and applications including: artistic compositions; aural histories; ubiquitous auditory displays; interactive sound installations; audio for different game platforms and sound for children's physical environments.

1 Introduction

“What a piece of work is a man, how noble in reason, how infinite in faculties, in form and moving how express and admirable, in action how like an angel, in apprehension how like a god.” William Shakespeare, Hamlet

As computation plays an ever larger role as an embedded part of the environment, research that seeks to understand the embodied nature of human's interactions with computation becomes increasingly important. *Embodied interaction* is an approach to understanding human-computer interaction that seeks to investigate and support the complex interplay of mind, body and environment in interaction. This approach is relevant for a broad class of ubiquitous computing systems that rely on direct physical interaction with computation embedded in the physical environment including mixed reality, tangible computing, responsive environments, pervasive computing and physical computing.

Dourish used the term *embodied interaction* to describe an approach to interaction design that placed an emphasis on understanding and incorporating our relationship with the world around us into the design and use of interactive systems (Dourish, 2001). Most simply, the approach involves leveraging users' natural body movements in direct interaction with spaces and everyday objects to control computational systems. Philosophically, the approach is based on a phenomenological paradigm that emphasises the role of action, perception and experience in meaning making. Knowledge is gained through purposeful and engaged practical actions in the world around us.

The term *intuitive* has been applied to such interfaces, although strictly speaking an interface cannot be intuitive since it is inanimate and cannot intuit anything. However, an interface can support people to intuit how to interact successfully with it, and it is that, this is meant by the term intuitive interface or *intuitive interaction*. Intuitive interaction is when a user can immediately use an interface successfully and the interface does what people expect (Spool, 2005). Many authors have suggested that movement of the body provides more natural or more intuitive form of interaction than the mouse, keyboard and screen set-up of a traditional desktop computer (e.g. Djajadiningrat et al., 2004; Hurtienne and Israel, 2007; Sharlin et al., 2004; Terrenghi et al., 2007; Ullmer and Ishii, 2000; Zuckerman et al., 2005).

This paper reports on a research project developed to better understand the cognitive mechanism that forms the basis for this claim. How and why might embodied interaction be an intuitive form of human-computer interaction? We suggest that the intuitive claim rests on specific aspects of the human body and the physical world that are drawn on, either implicitly or explicitly, in design. For example, one key aspect is the way that humans leverage embodied experiences through the use of metaphors to structure and understand abstract concepts. In this paper, we propose that an interface can be designed to support intuitive use by utilising physical body movements as the basis for interactional metaphors that relate to abstract representations. We present this design concept in the form of an intuitive interaction model for a responsive auditory environment. The interaction model uses an embodied metaphor to map body-based input with audio output composed of percussive¹ sounds. We further investigate the nature of intuitive interaction through an empirical experiment in which we compare 40 adult users' interactions using the intuitive, embodied metaphor interaction model versus a non-metaphor-based interaction model for the same responsive environment. In this paper, we report on the qualitative results of our experiment and describe how interactions are different for the two interaction models. We conclude by discussing how these results can inform embodied interaction design.

2 Background

2.1 Intuition (in apprehension how like a god)

Intuition (apprehension in Middle English) is a much maligned and often misinterpreted term. It has been used to refer to everything from a creative hunch to a spiritual insight. The Merriam-Webster online dictionary defines intuition as immediate apprehension or cognition; the power or faculty of attaining direct knowledge or cognition without evident rational thought and inference. Cognitive scientists have recently proposed that much of our thinking occurs not on stage, but off stage and out of sight (Myers, 2002). Thinking, memory and attitudes all operate on two levels: conscious and deliberate, and unconscious and automatic. Intuition is our capacity for direct knowledge, for immediate insight without observation or reason. It is thinking without conscious awareness. In contrast, deliberate thinking is reflective, reasoning-like, critical, analytic and operates in the realm of conscious awareness. Like the perceptual system, intuition operates through impressions. Judgements that directly reflect impressions are labelled intuitive (Kahneman, 2003).

Intuitive judgements are more or less accessible to an individual depending on a number of factors including: physical properties of the object of judgement (e.g. those which are routinely registered by the perceptual system such as size, distance, loudness and similarity); physical salience (e.g. larger items are more visually salient); priming through associative activation (i.e. contextual framing) as well as emotional and motivational states (Kahneman, 2003). Intuitive judgements can be overridden by a more deliberate, rational mode of operation. However, they may still affect subsequent responses through priming.

2.2 Embodied conceptual metaphors

We present embodied conceptual metaphors as a cognitive mechanism underlying intuitive interaction in direct interactional computational systems (e.g. responsive environments, gestural interfaces and tangibles). Intuitive judgements are unconsciously derived through embodied schema about appropriate movements and related embodied metaphors which link movements to system responses.

Rohrer (2006) describes a dozen different uses of the term embodiment. 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 (1980). We refer to these as *embodied conceptual metaphors* (or *embodied metaphors* for brevity). Embodied metaphors are an underlying mechanism behind instances of cognition that are often labelled intuitive. We explore Lakoff and Johnson's embodied metaphor theory with the aim of understanding a cognitive mechanism that designers may explicitly leverage to support intuitive interaction with embedded computation.

Lakoff and Johnson (1980) proposed a subclass of metaphor and a conceptual metaphor. They suggested that conceptual metaphors run deeper than simple linguistic conventions. Rather than just an interaction of two words, a conceptual 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 (1987) argued that metaphor is one of our primary cognitive structures for ordering experience. 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. Embodied schemata are mental representations of recurring dynamic patterns of bodily interactions that structure the way we understand the world. Conceptual metaphors extend embodied schemata to structure and organise abstract concepts. Embodied metaphors are based on embodied schemata and operate unconsciously and automatically.

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). Human experience a physical world in which sticks added to a pile or water added to a container result 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*. An orientational metaphor gives an abstract concept a spatial orientation. For example, *happy is up* and *sad is down*. This metaphor leads 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'. In this case, the use of orientational metaphors in understanding music is related to the emotional impact or content of the music.

In a responsive auditory environment an orientational metaphor could be used to relate sound parameters (e.g. volume, pitch) to spatial locations as exemplified by the English expressions 'high pitch' and 'turn up the volume'. This type of interface would likely be interpreted spatially based on specific musical instrument design conventions (e.g. high pitch is placed at the right of centre on a keyboard). To support interface

interpretation based on fundamental embodied schemata (rather than cultural conventions), we used an ontological metaphor based on body movement as described below.

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, categorised, grouped, quantified and qualified. Even when things are not discretely bounded, we refer to them in this way. For example, based on the metaphor music is a substance, we might say ‘The music flowed into the auditorium’. Alternatively, we can interpret music through the metaphor *music is physical body movement* (Barker, 1989; Jensenius, 2007). For example, ‘The music raced to its conclusion’. It is this metaphor that we focus on in this study. Other musical concepts, such as the principle kinds of musical process (e.g. melody, harmony and rhythm) and musical works themselves, are also often understood through spatial and physical metaphors but these are not the focus of our study.

2.3 *Metaphor and user interface design*

The use of metaphor in graphical user interface (GUI) design has been well covered in past research (e.g. Blackwell, 2006; Carroll and Thomas, 1982; Erickson, 1990; Laurel, 1986; Svanæs, 2001; Wozny, 1989). In GUIs, many elements of the interface are modelled 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 a classic example. Interface elements are modelled on common desktop objects (e.g. file folder and wastebasket) and actions on those objects (e.g. throwing out files into a wastebasket, opening files). While most HCI reference books include an entry for metaphor and give examples of the use of metaphor in user interface design, the potential benefits of using metaphor are not uncontroversial. As Blackwell (2006) points out, the understandings of the benefits and limitations of designing metaphor-based user interfaces have changed over time.

Most of the research, to date, on metaphor in user interface design has focused exclusively on the use of metaphor in the design of the visual communication elements of the GUI and in understanding user’s mental models of such interfaces. However, there are other ways in which metaphors may be used in interactive technology design including the use of metaphor in the interactional mappings between input actions and output responses. As computing becomes embedded in the physical environment and is enabled by direct body movements, gestures and physical object manipulations, it is critical to understand how metaphors may be used in interaction models. We conceptualise these interactional metaphors along the lines of Fishkin’s (2004) ‘metaphor as verb’. The use of metaphor to design visual and non-visual modalities of representation in movement-based interfaces will also be important. However, in this work, we focus on and isolate the aspect of interaction models involving the *mapping* between input actions and output responses.

While this mapping has been the subject of several recent papers (Antle, 2007; Koleva et al., 2003; Svanæs and Verplank, 2000), the focus on the use of metaphor in mapping has not been well studied. Metaphorical interaction models may be able to

support the user to intuitively enact appropriate input actions and understand the relationship to resulting system responses. They may also be instrumental in moving tangible and embedded interaction models beyond the limitations of direct mappings to deal with abstract concepts (Hornecker and Buur, 2006; Hurtienne and Israel, 2007).

2.4 *Intuitive interaction*

Intuitive interaction is one aspect of ease of use. One way in which intuitive interaction occurs is when, in a movement-based system, users enact appropriate input actions unconsciously or automatically, rather than consciously learning, step-by-step, how to interact with the system. Hurtienne and Israel (2007) proposed the following definition of intuitivity in the context of interaction with computation: “A technical system is intuitively usable if the users’ unconscious application of pre-existing knowledge leads to effective interaction”. Svanæs (2001) suggests that tangible and embedded systems can be made comprehensible to users in this way through the development of new interface metaphors. He suggests that for these interfaces to ‘make sense’ below the level of conscious awareness, they must be isomorphic with the basic structure of our experience of physicality. This does not necessarily imply a direct mapping that aims to duplicate reality. Rather, we suggest that it implies that designers of such mappings must be aware of how tacit knowledge of a physical source domain can be extended through conceptual metaphor to help users intuit how to interact with a more abstract conceptual target domain.

We explore the idea that conceptual metaphors, derived from embodied schemata and operating below the level of conscious awareness, may be used to create systematic and predictable relationships between specific human actions and specific system responses. We call these *embodied interactional models* and claim that they constitute a design principle which can be used to support intuitive interaction through the mechanism of how embodied metaphors structure human meaning making. While we expect some benefit to such designs, we do not expect that this benefit can be isolated from or made independent of other factors which impact the ease of use and quality of human-computer interaction with ubiquitous computing systems.

3 **The responsive audio environment**

To investigate the potential benefits and limitations of utilising an intuitive interactional model based on an embodied metaphor, we built several iterations of an interactive audio environment. In order to move beyond the predominantly visual paradigm of screen-based WIMP interaction, we explored the modality of sound based on a history of interactive musical composition systems. In order to move beyond the discrete actions or gestures commonly used in WIMP environments, we focused on the continuity of full body action in input. 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.

Interactive audio environments differ by the type of sensor used and their musical capabilities (Morales-Manzanares et al., 2001). We focus on system that uses camera vision to track full body continuous movement that is used to control percussive audio output. Winkler (1995) provides a good overview of movement sensing for interactive

music composition. Briefly, movement data may be mapped to musical parameters including volume, tempo, rhythm, pitch and beat. 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 explicitly be defined. Hunt et al. (2002) stated that by altering this mapping layer and keeping the interface itself and sound source constant, the entire nature of the environment (or instrument) are changed. Our implementation supports the inclusion of two different interaction models (mapping layers) in the same audio system to facilitate an experimental comparison.

3.1 *System inspiration*

Inspiration for our interactive audio environment came from Dalcroze Eurhythmics (Jaques-Dalcroze, 1972). In the 1930s, 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. Bodily knowing relies on tacit knowledge, the prereflective, automatic and unconscious knowing, we often label intuitive knowing. This reliance on and development of tacit knowledge in Eurhythmics is an appropriate domain in which to explore metaphors that support intuitive embodied interaction. For an excellent description of Dalcroze Eurhythmics (see Juntunen and Hyvönen, 2004).

3.2 *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 intuitively use 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 whole 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 and pitch). The primary criterion for the system was that the interface should be distributed in a space that facilitated movement (input) and produced variable sound responses (output). A second 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 maximise the potential to isolate embodied 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). This constraint was created to ensure that experimental results could be generalised beyond a population with a western music education.

3.3 *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 a pair of users' movements in a rectilinear space. The system relates locations, quantities and qualities of movement to changes in percussive audio output. Users control the instruments through their locations in the four quadrants of the space. Users control the sequencing of percussive sounds and the change of musical parameters of those sounds through the quantities and qualities of their collaborative body movements in the space (Figure 1). Videos of participants using The Sound Maker can be found at <http://www.antle.iat.sfu.ca/EmbodiedMetaphor>.

The Sound Maker environment was implemented using a Panasonic WV-CP470 3CCD camera, which fed video data through a Video Data DAC100 video digitiser 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 5.1×4.5 m 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 analysed the position data (as shown in Figure 2).

The system used sensed location data to infer users' speed, activity, proximity and flow. Speed was calculated for each user as the rate of change of user position in space. Activity refers to the amount of activity in users' movements (e.g. waving arms, stomping feet versus walking stiffly). 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 activity level is high and the speed is zero since speed is defined as the rate of change of position (in any direction). Proximity is the relative distance between the two users in the space. Flow is calculated from the frequency of speed changes of the users. For example, a series of rapid speed changes is characterised as a choppy movement. If each user is changing speed at different rates, the flow is also characterised as choppy.

Figure 1 Users moving around the space edges (left) and moving actively in one place (right) (see online version for colours)

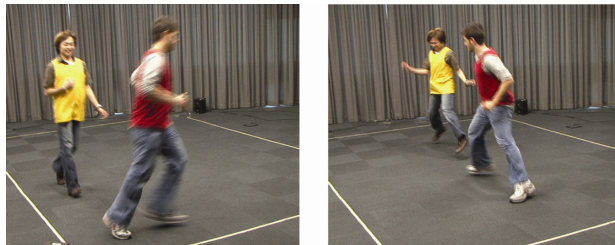
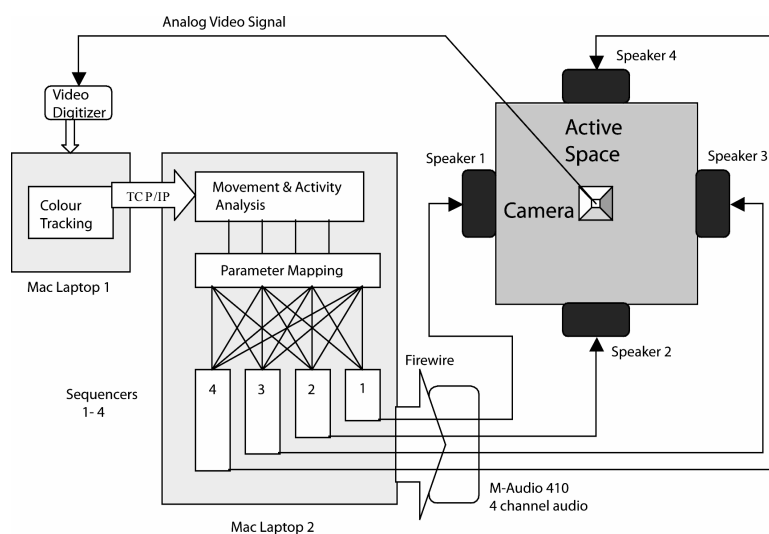


Figure 2 The Sound Maker system infrastructure

The location of users triggered specific sounds. Four percussive sounds (marimba, celesta, pizzicato viola and woodblock) were chosen to ensure performability of a wide variety of tempos and rhythms. The inferred movement characteristics were mapped to sound parameters: volume, tempo, pitch and rhythm. Volume varied continuously from loud to soft. Tempo varied continuously from very fast to slow. Pitch was discretised into categories to aid perception (high, medium, low – defined by frequencies). A rhythm is a recurring pattern formed by a series of notes differing in duration and stress. The rhythm was changed from steady through chaotic by introducing random variation in note duration for a specific event. The range from which the random change in note duration was chosen varied from a small range, resulting in a slightly degraded rhythm, to a large range, resulting in a much degraded, chaotic rhythm. Each of the four sound sequencers generated a rhythmic pattern based on the data being mapped to the control parameters and using an individually assigned sound. The output of each sequencer was then sent to one of four Yamaha biamplified monitor speakers placed along each side of the space.

3.4 Interaction models

The *music is (body) movement* metaphor suggests that music may be treated as human body movement. The human body movement source domain helps us to understand the musical sounds target domain. As suggested by Svanæs (2001), we focused on identifying embodied metaphorical mappings between source and target domains that preserved stable structural relations between the two domains.

In order to generate mappings based on this metaphor, we first deconstructed the metaphor in order to identify mappings that reflected shared structures between source and target domains. We then conducted a series of pilot studies based on early prototypes which instantiated these mappings. We looked for evidence of enactment or verbalisations based on our derived mappings. We created a set of potential mappings between input actions and output sound changes derived from our metaphor deconstruction and informed by our in pilot studies.

Table 1 Embodied metaphor-based mappings

<i>Movement</i>	<i>Parameter</i>	<i>Mappings</i>
Speed	Tempo	Fast is fast
		Slow is slow
Activity	Volume	More is loud
		Less is quiet
Proximity	Pitch	Nearer is high
		Farther is low
Flow	Rhythm	Smooth is rhythmic
		Choppy is chaotic

To avoid our own biases in validating the mappings, we conducted semi-structured interviews with four dancers and/or choreographers. These experts had undergone extensive training which included both experiential and reflective modes of understanding of the relations between body movement and music. We expected that they might be explicitly aware of and able to articulate metaphorical connections between movement and sound changes based on embodied knowledge. We expected that novice users would unconsciously enact these same mappings, but might not have been able to articulate them consistently. We first asked the experts to generate potential movement to sound parameter mappings. We then refined our original set of qualities and quantities of body movement based on results from these interviews and informed by Dalcroze Eurhythmics. For example, we constrained inputs to qualities of movement rather than specific types of movements (e.g. jumping and stomping) to avoid inconsistency in enactment. In a second interview, we asked the experts to match the refined set of input actions to changes in sound parameters. The results from these second interviews produced general agreement on mappings. 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. In determining our final mapping set, we eliminated movements that were difficult to sense (e.g. moving quietly). We also eliminated parameters for which there was no expert agreement (e.g. timbre).

We stress that the determination of mappings was non-trivial and required several iterations to arrive at the final mappings. Other mapping based on the *music is movement* metaphor are possible (e.g. speed to volume). However, we determined a set of four mappings which were consistent with the ontological metaphor, mutually consistent with expert opinion, excluded duplication and could be implemented in our camera-based system. The movement experts also validated the polarity of the mappings. Polarity refers to the direction of gradient of change. The final mappings for the intuitive interaction model based on an embodied metaphor are shown in Table 1.

Table 2 Non-embodied metaphor-based mappings

<i>Movement</i>	<i>Parameter</i>	<i>Mappings</i>
Flow	Tempo	Smooth is fast Choppy is slow
Proximity	Volume	Farther is quiet Nearer is loud
Speed	Pitch	Slow is high Fast is low
Activity	Rhythm	High is rhythmic Low is chaotic

The final mappings for the non-metaphorical model are shown in Table 2. We used the same movement variables and sound parameters as above. However, they were rearranged and the polarities reversed. None of these mappings were derived from the ontological metaphor *music is body movement*. In addition, neither users (in the pilot studies) nor experts enacted or identified these mapping consistently. As such, we labelled the model ‘non-intuitive’ meaning not based on the cognitive mechanism of embodied metaphor. While it is possible that a person might relate a characteristics of movement to a specific change in sound parameter based on a mapping in Table 2, it is likely that this interpretation would be a result of some other mechanism: a spatial metaphor, experience (e.g. proximity-volume), learning, chance or knowledge of acoustics (e.g. speed-pitch).

4 Methodology

4.1 Study design

In order to investigate, the potential benefits and limitations of utilising embodied metaphors, we designed an experimental comparison of interactions using an embodied metaphor based and non-embodied metaphor-based interaction model for the responsive musical environment involving the same tasks. We labelled the metaphor-based model ‘intuitive’ and the non-metaphor-based model ‘non-intuitive’ based on our theoretical foundations. We used a between-subjects design to eliminate any learning effects.

Recognising, mimicking and creating simple sound sequences with variations in volume, tempo, pitch and rhythm are common activities used to teach music (Juntunen and Hyvönen, 2004). Since participants were not required to have any musical training, beginner level exercises were chosen. Paired participants were asked to work together to create sound sequences by moving their bodies in the space. Sequences could be made with one or more of the percussive sounds. This type of movement-based exercise is common in the Dalcroze Eurhythmics approach to music education (1972).

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 sound parameter (volume, tempo, pitch and 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. The fifth and sixth tasks involved creating a sound sequence by varying two parameters at once (volume and tempo; pitch and rhythm). After users had had time to practice their sequence, they were asked to first perform the sequence, and then verbally explain it. After all structured tasks were completed, all participants were also given the opportunity to compose their own sound sequence, which they then demonstrated and explained to the facilitator.

4.2 Participants

Our final study was comprised of sessions with 20 pairs of adult volunteers (total 40 participants) of both genders (16 males and 24 females), aged predominantly in the 18–40 years range, recruited from the urban Simon Fraser University campus. No previous musical experience was required. All participants used computers daily. To support ecological validity and to reduce discomfort in a laboratory setting, we studied pairs of participants. Using pairs also increased the opportunity for verbal data collection. Participants were grouped randomly into pairs.

4.3 Measures

The tasks facilitated the collection of several forms of data, both quantitative and qualitative. The quantitative results and accompanying hypotheses are discussed in Antle et al. (in press). Our qualitative approach relied on detailed interactional analysis conducted both during the session and later through the review of video records. Intersubjectivity was achieved through the triangulation of three researchers who worked individually and then later together to identify salient interaction themes. In addition, session notes that focused on participant's task performance, interactional patterns and verbalisations were compared through close analysis to segments of video recorded material. Session notes included details of discrepancies between how participants performed (i.e. behaviours) and what they verbalised about their understanding of the system for each task. In addition, we recorded details of verbal explanations participants made about their experiences during tasks (spontaneously) and after each session in an informal interview.

5 Qualitative results

We observed that, overall, participants were more readily able to demonstrate and explain correct sound sequence patterns when using the 'intuitive' system. We summarise this finding by stating that, in general, the embodied metaphor-based system was easier to use than the non-embodied metaphor-based system. We noted some task dependency in this finding (discussed below). Further qualitative analysis of observational notes and recorded video material revealed five salient themes, which we use to characterise differences and similarities in how the participants interacted with the two interaction models.

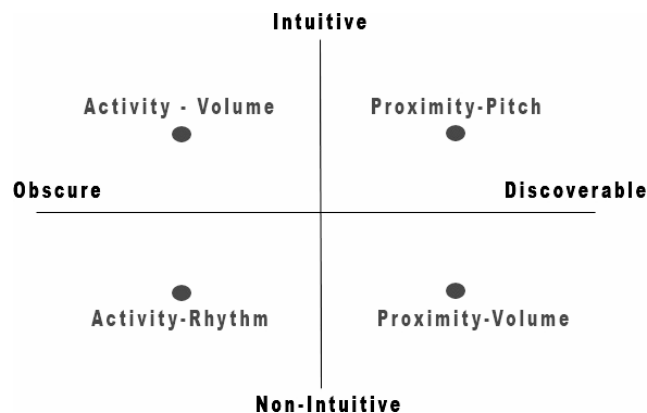
5.1 Intuitiveness and discoverability

We categorised our interaction models as intuitive and non-intuitive based on the presence or absence of an embodied metaphor in the mappings between input actions and system responses. While this characterisation is important, it alone could not predict participant responses. A second dimension was induced from observation. We propose that *discoverability* is a second dimension, which can be used to characterise the mappings in our interaction models and significantly influences participant interactions (see Figure 3).

Discoverability means how likely it is that participants will discover a mapping by chance. Discoverability, unlike intuition, also requires conscious reflection on the relationship between inputs and outputs. The mapping of proximity to volume in the non-intuitive interaction model proved to be highly discoverable. Many participants were able to accurately create the volume sequence using the non-intuitive system. When the participants moved far apart in space, the system volume decreased quickly to silence. As soon as participants 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. Participants' interpretation of this mapping may be related to acoustics – as a sound source moves away from a perceiver, the sound volume is perceived to decrease – however, no users verbalised this interpretation. The non-intuitive yet discoverable proximity to volume mapping appears in the lower right quadrant in Figure 3.

The inverse of discoverable is obscure. A mapping is obscure if it is unlikely that it will be revealed by chance actions accompanied by salient feedback. For example, when participants used the activity level to rhythmic sounds mapping in the non-intuitive version they verbalised that it was difficult to listen while moving actively. This made it difficult to discover that rhythmic sounds were associated with high levels of activity. High levels of activity felt chaotic and thus sounded chaotic. The activity level to rhythm mapping appears in the lower left quadrant in Figure 3. It is both non-intuitive and obscure. Based on quantitative accuracy results reported in Antle et al. (in press), the mappings of speed to tempo and flow to rhythm can be characterised as both intuitive and discoverable. The mappings of speed to pitch and flow to tempo are both non-intuitive and obscure. The remaining mappings are characterised as shown in Figure 3.

Figure 3 Mappings: intuitive versus discoverable



5.2 *Experiential and reflective modes of cognition*

There are many modes of cognition. Experiential cognition involves a state in which we act, perceive and react to the world around us effortlessly and efficiently (Norman, 1993). What we have defined as intuitive interaction can be classified as a kind of experiential cognition. Reflective cognition involves a state in which we consciously compare, contrast and make decisions (Norman, 1993). We expected that participants would view the floor space as a physical one in which experiential, unconscious, action-driven cognition would hold a privileged role in interaction. Instead, many participants tended to approach the space as a computational one, which required reflective, analytical thought in order to understand how it worked. This may be an artefact of the user population as all participants used computers on a daily basis. However, an analysis of the chronology of sessions revealed a common pattern of experiential and reflective approaches to interaction.

Participants' first explorations of both systems tended to involve effortlessly moving around the floor space varying many of the quantities and qualities of movement we had identified as intuitive and built into the system. Action appeared to drive their explorations. We characterise this as an experiential, prereflective driven approach to using the system. We observed that when participants became unsure about the system's response or misinterpreted a correct response as an incorrect one (i.e. a breakdown occurred), they quickly changed to a more analytical approach in which they systematically tested different movement sequences. For example, we observed many instances when a pair would communicate with each other to make a plan for action, execute the plan and then reflect on the results. We characterise this as a conscious, reflective, deliberate and reasoning driven approach to using the system. We observed that near the end of tasks, if participants had not yet found a solution and were given the final 1-min warning, they often resorted back to the experiential approach, moving around the space in ways that we had envisioned and labelled intuitive.

Distinguishing human cognition into two discrete categories is simplistic. However, we can characterise cohesive segments of participant's task solving activity as either predominantly experiential or reflective. The proportion of session time spent predominantly in each mode can be used to distinguish interaction patterns between the intuitive and non-intuitive systems. In both systems, we observed that adult participants spent more time in a reflective mode than in an experiential mode. We also observed more and longer instances of reflective, analytical approaches in sessions involving the non-intuitive system. These sessions were typically characterised by fewer instances of successful demonstration or verbalisation of sound sequences for all tasks. We conclude that the embodied metaphor-based system facilitated a proportional mix of experiential and reflective interaction that resulted in users more successfully learning to control and understand the system. However, we stress that for users to fully understand how to create particular effects required that they utilised both experiential and reflective cognition.

While it is possible to characterise a portion of a task session as predominantly experiential or reflective, a finer level of interactional analysis revealed that the two modes overlap and often operate simultaneously. Post-session debriefings revealed that while input actions may be enacted intuitively, interpretation of resulting sound feedback was often processed consciously with an intention to solve the task. For example, participants stated that they 'just moved around' and while they were moving 'thought

about what was happening'. That is, learning how to control the sound parameters required that participants recognised when their movements elicited the sound effect they were trying to create. It also required that they remember how they were moving to create a particular sound change and be able to duplicate it. For example, one participant said her partner, "That's it (the sound effect), what did I just do?" Behavioural observations cannot discern if certain mental activities occurred consciously or not. However, it seems likely that a complex interplay of experiential knowing and reflective analysis was required to successfully complete sound sequence tasks. We suggest that an interplay of intuitive and reasoning-based operations was evident, as described by Kahneman (2003). Intuition and perception may have worked together to generate automatically processed impressions. Or intuition may have been largely overridden by reasoning and conscious interpretation of sounds. While it was not observable, theory suggests that intuitive impressions may still have an effect by priming reasoning.

We observed more instances of experiential cognition in the embodied metaphor-based system and we associate this, in part, with successful task performance. While intuition, or unconscious application of preexisting knowledge, may drive initial and time pressured interactions, this approach was interspersed with analytically driven interactions. Neither has a privileged role. The two approaches work together. When one fails, the other is called on, perhaps unconsciously. We can conclude that in the embodied metaphor-based system, users' intuitions based on unconscious embodied knowledge supported appropriate interaction and correct system interpretation.

5.3 *Oriental versus ontological structures*

We observed in the pilot studies and in this study that users' exploratory behaviours were correlated to the shape of the active sensing environment. Despite directing participants to consider how they moved in the space, many participants routinely explored the spatial aspects of the environment first. In particular, they were fascinated by edges (Figure 2 (left)) and the locations of the speakers. They routinely moved linearly (Figure 2 (right)). Many of their verbalisations showed a focus on where the sounds were located or triggered ('Here its low') rather than focusing on how to make sounds ('When I move this way, it gets lower'). Their verbalisation less frequently concerned references to the quantity or quality of their movements. In general, we observed spatial exploratory behaviours and noted spatial references in their verbalisations more frequently in the non-intuitive system. We also observed more instances of spatial explorations in the latter portion of sessions with the non-intuitive system. In the post-session debriefing, participants in both groups described a mental model of the system in spatial rather movement terms. This corresponded to some of their behaviours during the tasks. We conclude that, in general, orientational rather than ontological metaphors dominated users' exploratory actions and mental models of the system. This spatial interpretation was more evident in interactions with the non-intuitive system.

5.4 *System artefacts*

Various authors developing ubiquitous computing systems have noted that users tend to exploit the limitations of a system to their advantage. For example, Benford et al. (2006) found that players of the game 'Can you see me now?' often exploited GPS blackspots (e.g. building shadows) to hide from other players. In our study, we also observed this

behaviour. System artefacts sometimes worked to participants' advantage and sometimes to their disadvantage. System latency in sound display combined with some motion tracking aberrations resulted in some inconsistent feedback. Participants sometimes discovered these artefacts and used them to create patterns. For example, when participants changed their motion just as the system was responding to it, they commonly interpreted the response as confirmation that the new motion produced the effect. Some participants tested out their theory while others claimed they had found the solution and were ready to demonstrate it. However, when they went to make the correct pattern they could not always replicate it reliably.

In some cases, these artefacts were stable and participants reproduced them successfully. For example, in the intuitive system when a participant moved rapidly towards the edge of the space, reached the edge, stopped and moved rapidly in the opposite direction, they often described their movement as moving quickly (ignoring the instance of stopping to turn around). However, their movements were interpreted by the system as a sequence of fast speed, zero speed and fast speed, and they were able to create a tempo sound sequence of fast-slow-fast. One participant described this solution as 'moving fast towards an edge' rather than moving quickly, slowly and then quickly. In these kinds of cases, the prototype afforded participants a variety of ways in which to solve and interpret their solutions to the tasks. Ultimately, it is the interpretation of the system's responses by a user rather than technical parameters that determines her understandings.

5.5 *Personal factors*

Analysis of task completion strategies revealed two distinct groups of users. We characterised these two groups as perfectionists versus users who tolerated ambiguity. Perfectionists often tested and retested their solutions before indicating they were ready to demonstrate. Ambiguity tolerant pairs found a solution or partial solution and immediately wanted to demonstrate it without further testing. This distinction highlights that users may differ in their willingness to approach the system experientially relying on intuition rather than analytically relying on experimentation.

In addition, different degrees of personal agency were seen in paired interactions. A common pattern was one in which participants would explore by themselves, and then explain what they found to their partner; even after being told the space was meant to be used collaboratively. Dividing tasks between partners was also a common strategy to experimentation. For example, one participant said, "You find a slow one and I'll find a fast one". Individual approaches did not produce the desired pattern since input actions were designed to be cumulative. It is unclear how paired interactions may have affected task completion strategies. This warrants further investigation.

6 **Summary and implications**

The comparison of an embodied metaphor- and non-embodied metaphor-based interaction model for a responsive audio environment allowed us to observe and reflect on fundamental differences between interacting with a system where one version included a mapping layer based on a body-based source domain and sound-based target domain and one that did not. We found evidence that including an embodied interactional

metaphor in an interactive environment could support intuitive interaction through the cognitive mechanism of embodied schema extended through metaphor to explicate abstract (sound) representations. Users' unconscious application of preexisting knowledge in the form of embodied knowledge about relations between movements and sound parameter changes led to effective interaction. However, participants often interacted with the system in ways that were more conscious and reflective than intuitive, and thus, they did not always leverage the embodied metaphor-based mappings between movement and sound response.

We observed participants interacting with both versions in shared and unexpected ways, which have significant design implications. Firstly, and perhaps most significantly, we propose the characterisation of a mapping as intuitive or non-intuitive (embodied metaphor based or not) is not sufficient to determine if it can be easily intuited. In order for a mapping to be intuitively enacted, it must also be discoverable. This highlights the role of perceivable feedback in supporting intuitive input actions. Intuition is more or less accessible depending on salience, context and how easily precepts are registered by the perceptual system (Kahneman, 2003). Further, mappings that are easily discoverable but not intuitive also can be readily and quickly learned if and when they are stumbled upon by chance and if resulting feedback is unambiguous. The distinction between aspects of interaction that are intuitive and those that can be immediately learned is blurry.

Secondly, we observed that many participants predominantly approached learning to use both systems not through an experiential, intuitive, action-driven approach but through an analytical and deductive approach. However, we did observe more segments of experiential active-driven cognition in the metaphor-based system. Designers should consider how they might promote an appropriate mixture of experiential and reflective approaches to system use for interactive environments. This consideration must be based on intended system usage as noted by Norman (1993). For example, in an interactive audio environment intended for music composition or performance, a predominantly experiential approach should be supported. We suggest that the inclusion of embodied metaphor in the interactional model can help support a predominantly experiential approach to interaction. If intuitive actions result in desired system responses, then reasoning about how the system might work is not required. Conversely, in an environment intended to teach concepts related to music or sound, support for the kind of reflective thought required for knowledge acquisition should be considered. In this case, the non-intuitive mappings might encourage reflection on underlying concepts. Non-intuitive yet discoverable mappings can be used to support reflection. This approach has been demonstrated in sensor-based interactive learning applications in which pairing of familiar actions with unfamiliar responses was utilised to support reflection (e.g. Rogers and Muller, 2005).

In order to explore this design issue further, we have executed a similar study with forty children aged 7–10 years old. Children were found to utilise an experiential approach to learning and rely on embodied knowledge more than adults (Antle et al., 2008). However, while children using the metaphor-based system learned to control the system more accurately than those in the non-metaphor-based system, it was unclear if they explicitly learned more about sound concepts. Thus, the nature of intended use and the characteristics of the end-user group must both be considered in determining an optimal mixture of support for experiential and reflective interaction.

Finally, the predisposition of participants to explore and interpret both system versions spatially leads us to propose that orientational metaphors rather than ontological ones dominated intuitive embodied knowledge and mental system models in the context of using this kind of responsive environment. While this premise needs further exploration, we suggest that distinguishing between orientational and ontological metaphors in embodied interaction design is critical. We also suggest that using ontological metaphors may have wide utility but that they need to be used in ways that distinguish them from orientational metaphors.

7 Conclusions and future work

This paper reports on an exploratory study in which we compared two interaction models for the same system in order to explore the idea that incorporating an embodied metaphor in the interaction model may make a system more intuitive to use. We found evidence that an embodied metaphor could be successfully used to create systematic and predictable relationships between specific human actions and specific system audio responses. Qualitative analysis revealed that the system was easy to use if it instantiated both intuitive mappings between input actions and output responses (through the mechanism of embodied metaphor) and if these mappings could be easily discovered through chance actions in the space accompanied by perceivable system feedback.

The project described in this paper contributes to understanding the role and the potential benefits that leveraging embodied metaphors in design may have for ubiquitous computing systems that rely on direct physical interaction with computation. We conclude that in apprehension man may be like a god, however, designers who understand the cognitive basis for such apprehension may be able to leverage these cognitive mechanisms in their designs. These pursuits form an active and important research area in ubiquitous computing. It is critical that we empirically examine claims of the potential benefits of embodied interaction and in doing so seek to understand the cognitive basis for such claims. This knowledge contributes to understanding the challenges and limitations of designing and evaluating interfaces that explicitly support embodied interaction and contributes to the maturation of the field.

Acknowledgements

The authors wish to thank the anonymous reviewers for their thoughtful feedback. This research was funded by the National Science and Engineering Research Council of Canada (NSERC). Special thanks to all the SFU students who participated in this study.

References

- Antle, A.N. (2007) ‘The CTI framework: informing the design of tangible systems for children’, *Proceedings of Conference on Tangible and Embedded Interaction*, Baton Rouge, LA, USA: ACM Press, pp.195–202.
- Antle, A.N., Corness, G. and Droumeva, M. (in press) ‘What the body knows: exploring the benefits of embodied metaphors in hybrid physical digital environments’, *Interacting with Computers: Special Issue on Physicality*, Elsevier.

- Antle, A.N., Droumeva, M. and Corness, G. (2008) 'Playing with the sound maker: do embodied metaphors help children learn?' *Proceedings of the Interaction Design for Children*, Chicago, IL, USA: ACM Press, pp.178–185.
- Barker, A. (1989) *Greek Musical Writings: Volume II, Harmonic and Acoustic Theory*, Cambridge, UK: Cambridge University Press.
- Benford, S., Crabtree, A., Flintham, M., Drozd, A., Anastasi, R., Paxton, M., N., T., Adams, M. and J., R-F. (2006) 'Can you see me now?', *ACM Transactions on Computer-Human Interaction (TOCHI)*, Vol. 13, pp.100–133.
- Blackwell, A. (2006) 'The reification of metaphor as a design tool', *ACM Transactions on Computer-Human Interaction (TOCHI)*, Vol. 13, No. 4, pp.490–530.
- Budd, M. (2003) 'Musical movement and aesthetic metaphors', *British Journal of Aesthetics*, Vol. 43, No. 3, pp.209–223.
- Carroll, J. and Thomas, J.C. (1982) 'Metaphors and the cognitive representation of computing systems', *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 12, No. 2, pp.107–116.
- Djajadiningrat, T., Wensveen, S., Frens, J. and Overbeeke, K. (2004) 'Tangible products: redressing the balance between appearance and action', *Personal and Ubiquitous Computing*, Vol. 8, No. 5, pp.294–309.
- Dourish, P. (2001) *Where the Action Is*, Cambridge, MA: The MIT Press.
- Erickson, T. (1990) 'Working with interface metaphors', in L. Brenda (Ed.) *The Art of Human-Computer-Interface Design*, New York, NY: Addison-Wesley.
- Fishkin, K. (2004) 'A taxonomy for and analysis of tangible interfaces', *Personal and Ubiquitous Computing*, Vol. 8, No. 5, pp.347–358.
- Hornecker, E. and Buur, J. (2006) 'Getting a grip on tangible interaction: a framework on physical space and social interaction', *Proceedings of the Human Factors in Computing Systems*, Montréal, Québec, Canada: ACM Press, pp.437–446.
- Hunt, A., Wanderley, M. and Paradis, M. (2002) 'The importance of parameter mapping in electronic instrument design', *Proceedings of the Conference on New Instruments for Musical Expression*, Dublin, Ireland, Singapore: National University of Singapore Press, pp.1–6.
- Hurtienne, J. and Israel, J. (2007) 'Image schemas and their metaphorical extensions: intuitive patterns for tangible interaction', *Proceedings of the Tangible and Embedded Interaction*, Baton Rouge, LA: ACM Press, pp.127–134.
- Jaques-Dalcroze, E. (1972) *Eurhythmics Art and Education*, New York, NY: Benjamin Blom Inc.
- Jensenius, A.R. (2007) 'PhD thesis: action-sound: developing methods and tools to study music-related body movement', Available at: <http://www.arj.no/research/>.
- Johnson, M. (1987) 'The body in the mind: the bodily basis of meaning, imagination, and reason', Chicago, IL: Chicago Press.
- Juntunen, M. and Hyvönen, L. (2004) 'Embodiment in musical knowing: how body movement facilitates learning within Dalcroze Eurhythmics', *British Journal of Music Education*, Vol. 21, No. 2, pp.199–214.
- Kahneman, D. (2003) 'Nobel prize 2002 lecture: maps of bounded rationality: a perspective on intuitive judgment and choice', in Frängsmyr (Ed.) *Les Prix Nobel*, Stockholm: Nobel Foundation.
- Koleva, B., Benford, S., Ng, K. and Rodden, T. (2003) 'A framework for tangible user interfaces', *Proceedings of the Mobile HCI*, Udine, Italy, pp.46–50.
- Lakoff, G. and Johnson, M. (1980) *Metaphors We Live By*, Chicago, IL: Chicago Press.
- Laurel, B. (1986) 'Interface as mimesis', in D.A. Norman and S.W. Draper (Eds.) *User Centered Systems Design*, Hillsdale, NJ: Lawrence-Earlbaum.
- Morales-Manzanares, R., Morales, E., Dannenberg, R.F. and Berger, J.F. (2001) 'SICIB: an interactive music composition system using body movements', *Computer Music Journal*, Vol. 25, No. 2, pp.25–36.
- Myers, D.G. (2002) *Intuition: Its Powers and Perils*, USA: Yale University Press.

- Norman, D.A. (1993) *Things That Make Us Smart*, New York, NY: Basic Books.
- Rogers, Y. and Muller, H. (2005) 'A framework for designing sensor-based interactions to promote exploration and reflection', *Int. J. Human-Computer Studies*, Vol. 64, No. 1, pp.1-15.
- Rohrer, T. (2006) 'The body in space: dimensions of embodiment', in J. Zlatev, T. Ziemke, R. Frank and R. Dirven (Eds.) *Body, Language and Mind*, Vol. 2, Berlin, Germany: Mouton de Gruyter.
- Sharlin, E., Watson, B., Kitamura, Y., Kishino, F. and Itoh, Y. (2004) 'On tangible user interfaces, humans and spatiality', *Personal and Ubiquitous Computing*, Vol. 8, No. 5, pp.338-346.
- Spool, J. (2005) 'What makes a design seem intuitive?' *User Interface Engineering*, Available at: http://www.ue.com/articles/design_intuitive/.
- Svanæs, D. (2001) 'Context-aware technology: a phenomenological perspective', *Human-Computer Interaction*, Vol. 16, No. 2, pp.379-400.
- Svanæs, D. and Verplank, W. (2000) 'In search of metaphors for tangible user interfaces', *Proceedings of the Designing Augmented Reality Environments*, Elsinore, Denmark: ACM Press, pp.121-129.
- Terrenghi, L., Kirk, D., Sellen, A. and Izadi, S. (2007) 'Affordances for manipulation of physical versus digital media on interactive surfaces', *Proceedings of the Conference on Human Factors in Computing Systems*, San Jose, CA: ACM Press, pp.1157-1166.
- Ullmer, B. and Ishii, H. (2000) 'Emerging frameworks for tangible user interfaces', *IBM Systems Journal*, Vol. 39, Nos. 3/4.
- Winkler, T. (1995) 'Making motion musical: gesture mapping strategies for interactive computer music', *Proceedings of the Computer Music Conference*, San Francisco, CA: Computer Music Association, pp.261-264.
- Wozny, L.A. (1989) 'The application of metaphor, analogy and conceptual models in computer systems', *Interacting with Computers*, Vol. 1, No. 3, pp.273-283.
- Zuckerman, O., Arida, S. and Resnick, M. (2005) 'Extending tangible interfaces for education: digital Montessori-inspired manipulatives', *Proceedings of the Conference on Human Factors in Computing Systems*, Portland, Oregon: ACM Press, pp.859-868.

Note

¹A percussion instrument is any object which produces a sound by being hit with an implement, shaken, rubbed, scraped or by any other action which sets the object into vibration. Examples include: drum, xylophone and cymbals.