Vuzik: the Effect of Large Gesture Interaction on Children's Creative Musical Expression

Junko Ichino
Kagawa University, Japan
ichino@eng.kagawa-u.ac.jp

Aura Pon
University of Calgary, Canada
aapon@ucalgary.ca

Ehud Sharlin
University of Calgary, Canada
ehud@cpsc.ucalgary.ca

David Eagle
University of Calgary, Canada
eagle@ucalgary.ca

Sheelagh Carpendale
University of Calgary, Canada
sheelagh@cpsc.ucalgary.ca

ABSTRACT
Bringing the body more fully into interaction has attracted attention as research now looks to combines body, mind, cognition and emotion when people interact with a digital environment. This paper describes a user study comparing *Vuzik* – an application where the user manipulates, arranges, and composes music with painting interaction akin to that used when standing at an easel, to a traditional, GUI based musical interface. *Vuzik* aims to promote creative musical experiences in children by allowing the child's actions and movements as he/she paints on a large display resembling a canvas using a palette and brush or finger to control musical parameters interactively. In a study conducted with fourteen elementary school children, we found that when compared to a more WIMP-based traditional tool, *Vuzik* promoted larger scale gestures, ease of learning, and the formation of a broader overall understanding of their musical creation.

Author Keywords
Embodiment; large gesture interaction; musical expression; children, creativity.

ACM Classification Keywords
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *Evaluation/methodology, Interaction styles*; J.5 [**Arts and Humanities**]: *Performing arts.*

INTRODUCTION
For children, singing or creating sounds is a mental activity that can be essential for life and growth. At the same time, music is an indispensable part of forming an image of the world by exploring the causal relationship between one's own actions and changes in the environment [32]. Many past efforts attempted to harness automation and technology and apply them to artistic and creative context, including music. However, such efforts sometimes indicated that tools which use advanced technology, far from sufficiently stimulating human creativity and sensibility, instead inhibit it (e.g. [2]). When holding a mouse or a keyboard, which enable precise and detailed information input, and looking at detailed and ordered objects shown on a display of at most 20 to 30 inches, people become engrossed in the detail of it, getting caught up in the simple operation of these devices. As a result, an important consideration, that of developing an overall mental image of the desired creation and making bold alterations to the creative work, can be lost. This is said to lead to design which is superficially pretty but lacking in content and substance [31].

OzCHI '14, December 02 - 05 2014, Sydney, NSW, Australia

http://dx.doi.org/10.1145/2686612.2686649

On the other hand, embodied interaction is attracting attention as a field that combines body, mind, cognition and emotion when people interact with a digital environment [8]. Dourish states in *Where the Action Is* that while our bodies are our most familiar presence, it is a presence which is difficult to consider objectively [7]. Consequently, interaction design has been mostly based on the Cartesian principle of mind-body dichotomy. He stated that future interaction design must give careful thought to the user's Heideggerian state of immersion in the everyday world – embodiment – and that digital technology must be used properly [7]. In other words, what embodiment points at is not only the body itself, but something abundantly suggestive of sensibilities, senses and intuition. Following, in a digital environment that was designed with embodiment in mind, user experience will rely on rich mappings between physicality, body gestures and movements, tangible artifacts, and the interface, resulting in experiences that relate directly to users' senses and actions.

In this paper we explore an interface, *Vuzik*, which maps large gesture interaction using tangibles in front of a large interactive screen to creative musical experiences [12, 22]. Music can inspire and lead to a whole body reaction and movement, naturally, as seen in dance. Conversely, musical instruments are devices that convert body movements into sound. In this sense the relationships and mappings between our bodies and music are important elements when creating interactive media that should musically engage people. Attempts to merge more holistic interaction systems and creative expression also coincide with a growing interest in the relationships between interaction design, body movement [16], and artistic expression [21].

In this paper we propose an interactive system aimed at children – *Vuzik* – to express music that maps larger-scale gesture interaction to visual and auditory feedback. We closely study the design space, the possible mappings of physicality and gestures to musical expression, and compare *Vuzik* to a more traditional, GUI-based musical interface. We expect that *Vuzik* will contribute to the design of interactive systems that enable and foster creative expression.

RELATED WORK

Interfaces Supporting Children's Creative Activities

Many systems and interfaces were designed to support children's creative activities and expression in the areas of music, art, and storytelling.

There are many interfaces that aim at allowing children to create and express themselves musically. For example, Hyperscore [9], Making Music [30] and Dragencer [13] allow the user to compose music by freehand line drawing. In Hyperscore, many notes are entered in one sequence by drawing lines with the mouse, with detailed adjustments are made at a later point. FlexiMusic Kids Composer [10] are authoring tools for creating music by combining recorded sound and prepared pieces of music. Singing Fingers [25] is a drawing and sound sampling tool for the iPad. Speech and surrounding sound is recorded while lines are drawn with a finger, later these lines can be traced allowing the recorded sound to be played back.

There are also many interfaces that support children's drawing. For example, Jabberstamp [23] allows pictures to be drawn onto paper are placed on a tablet, and sound can recorded when specific places in the picture are touched with a stamp or trumpet with an embedded stylus. When these locations are later traced with the pen, the recorded sound is played back. IO Brush [26], traces objects in the real world using a brush with an embedded camera, sampling the color and movement of the object. These samples become drawing materials to make pictures on a canvas using real world colors and patterns. The Sound of Touch [15] takes the form of a hand-held wand that contains a microphone, piezo vibration sensor, and pushbutton. The wand can be used to record sound and later play back the recording by brushing, scraping, striking or otherwise physically manipulating the wand against physical objects.

Examples of studies focusing on creating stories are TellTable [5], StoryMat [27] or StoryRoom [17]. TellTable is a system where children create and share stories on a table top display. Children create characters and a background by taking pictures of real world objects or people with a capturing tool and by drawing pictures on the table. By placing these on the table and recording and adding story narration or dialogue, a complete story is created.

Gestural Interaction for Musical Expression

The number of studies investigating interface design for musical expression using large-scale gestural interaction is gradually increasing with the development of sensor technology.

LaserWall [20] can track hand and body motion in front of large interactive media surfaces. BodySpace [29] can control a music player by using feedback from acceleration sensors attached to various places on the body such as the head, chest and the hips. With "The Bow is Bent and Drawn" [33], musical pieces which are rich in social and active expressions can be both listened to, and created. While not musical in the traditional sense, Williams et al.'s evaluation of SignalPlay [34] describes how participants' manipulation of augmented physical objects enabled them to construct, interpret, and manipulate connections between a physical space and a correlated soundscape. For BodyBeats [35], three types of whole body interaction systems were developed – jumping on a trampoline, touching a wall-sized push-pad with several people, and changing hand position or posture – and the authors explore the possibilities for whole body movement as an interface for children to generate music. MINWii [1] is a musical therapy game aimed at Alzheimer's patients where virtual keyboards are played using a Wiimote or where music that has been prepared beforehand is performed. One Man Band [3] is a musical game using a Wiimote in a similar way, where music is performed by controlling the timing and the type of instrument played using gestures. Articulated Paint [14] enables users who are not music experts to control a piece using gestures, moving a brush like a conductor's baton. When drawing lines with a brush containing an embedded pressure-sensitive sensor on the display, parameters to enrich the music's expressive power, such as the piece's tempo and dynamics, articulation or vibrato, are determined from the brush stroke information. In Moboogie [11], the three axes of the acceleration sensor embedded in an Android mobile device are mapped to melody, base and drum, respectively, and allows interactively altering musical pieces that were provided beforehand according to the user's movements.

These systems, primarily allowing manipulation of pre-created music, provide an entertainment space where music is explored using large gesture interaction.

Positioning of This Effort

Vuzik is an interactive system that enables a user to experience creative musical expression while while painting with the same large body gestures that can be used at an easel This shares many elements with the systems described above: Hyperscore, Making Music and Dragencer are similar to *Vuzik* in that they support children's expression and creation of music through drawing actions. However, they emphasize small mouse and finger interactions instead of our larger expressive gestures. BodyBeats and Moboogie share *Vuzik*'s aim of supporting musical expression using one's body, but as they are not aimed at creativity in the sense of allowing users to create and compose new music, these projects' goal differs from *Vuzik*'s.

VUZIK

Design Concept

We established the following design concepts for the design of *Vuzik*'s functionality and interaction to support creative musical expression for children.

Large gestural movement: Construct a space for interaction with the system not only through fingers or hands, but through larger movements.

Keeping in touch with the physical: Promote the use of physical artifacts for children to explore, discover and play with in order to easily acquire embodiment.

External representation: One of the aims of *Vuzik* is to understand the role of larger, fuller gestural movement in creative activities. An interface where the relationship between music and embodiment is more explicitly constructed is therefore important.

Free expression: Ensure that children can express their imagination as freely as possible to encourage creativity. A certain amount of molding and structuring of the user's movement through the system design is unavoidable, but not using any music piece provided beforehand minimizes the restrictions or influence on the child's own creativity.

Playfulness, simplicity, and immediacy: As *Vuzik* intended users are children, emphasize playfulness and simplicity in designing both the software user interface and the physical tangible components. Also, ensure that children can create sound instantaneously without the need for elaborate planning.

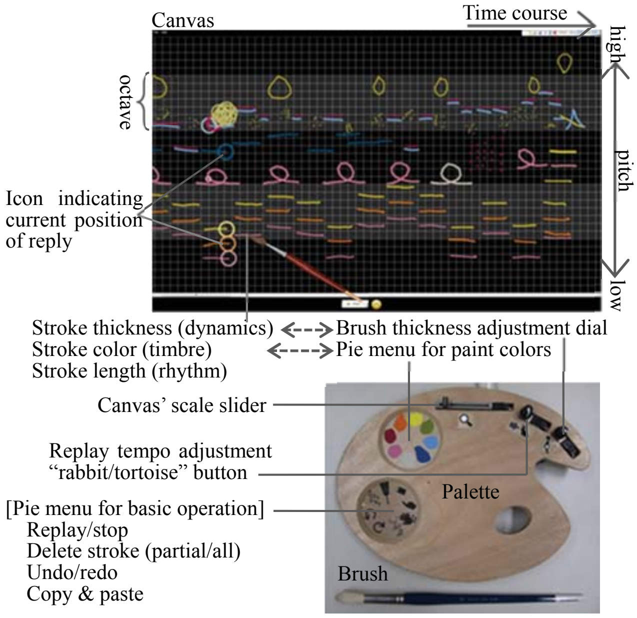


Figure 1. Basic musical elements of *Vuzik* and their mapping.

System Description

We employed an iterative and incremental design process when developing *Vuzik*, incorporating continuous evaluation of our prototypes by children participants.

"Drawing" Music Using Larger Scale Gestures

In accordance with our design concepts, this study employed a classic yet more intuitive brush interface style of drawing a picture on a large-sized canvas using a finger or a brush, where the correspondence between body movement and manipulation is clear. Because the action of drawing a picture on a large canvas requires the same kind of larger gestural interaction as when playing a general musical instrument, we ensure this study focuses on body expressions using much more of the body than previous interfaces. Also, sketching by hand is regarded as an effective means of externalization for the creative thinking process [18].

Mapping to Basic Musical Elements

Music is composed of the four basic elements of rhythm, pitch, dynamics and timbre. To satisfy our design concept we mapped each element to the interface (Figure 1).

Canvas: On the canvas, the horizontal axis represents time, where the music progresses from left to right, and the vertical axis represents the pitch, where the sound becomes higher as it moves upwards from the bottom. The complete image drawn onto the canvas represents one piece of music, starting at the left edge and ending at the right edge.

To investigate the way in which children "draw" music, we carried out an informal user study. When we asked children to express a melody they had heard in a drawing, we were able to divide their styles of expression into two broad types, namely the free style (i.e., scribbles, letters, dots, pictures of hands) and the matrix style (i.e., graphic depiction of the perceived musical performance, including depictions showing an awareness of meter including rests). In order to deal with both styles of expression, in the background of the canvas we prepared a mode where additional lines were shown as a reference for the pitch and the time axes, as well as a mode that did not show this information.

To enable simple and smooth movement between viewing an overview of the complete piece and focusing on a detailed part while creating the music, we ensured that, instead of creating sub- windows, one and the same window (canvas) can be displayed in any scale through zooming. The user can change the scale of the canvas using the scale slider on the palette.

Brush, finger: One stroke on the canvas using either a finger or a brush consists of one or more notes connected smoothly through a slur. The pitch of each note is mapped to the value on the vertical (y-axis). The positions on the y-axis corresponding to the points within the stroke represent the changes in pitch, and the time from the previous changing point to this one provides the rhythm. The timing of the start of playing a stroke on the canvas is mapped to the values on the horizontal axis (x-axis) formed by the points at the start of the stroke. The dynamics correlate to the thickness of the brush stroke and the timbre to the paint colors selected. The timbres are mapped to eight colors of paint by one of the authors, a professional composer and oboe player, generated timbres aiming to analogously match the character of the colors.

Palette: The palette is operated with finger touch. Almost all the application's manipulations are performed on the physical palette rather than the GUI. This includes the following: event operation, to start and stop the replay of created music; changing the properties of the brush, i.e., the thickness (dynamics) and the paint colors (timbre); regulating tempo and volume of the reply; zooming on the canvas; and basic operations including copy or undo strokes (notes). For simplicity, we kept the number of functions to a minimum. We designed the system so that the children could move freely within a wide area holding the palette like a painter in their non-dominant hand, or put it next to the canvas and move around empty-handed.

Two Interaction Modes

This system is composed of two interaction modes: the Creation mode, to create music, and the Playback mode, to playback and listen to the music. The children can move between them seamlessly without regard to two modes.

Creation mode: When the user, using a brush or their fingers, touches the display canvas directly, corresponding sound is played back simultaneously with a line being drawn freehand on the display. The brush properties can be changed on the palette. The user can change the paint color (timbre) by touching a finger to the color that mimics paint on the palette, and they can change the thickness of the brush (dynamics) by turning the dial on the palette.

Playback mode: When the user touches the "note" icon on the palette, the music is played back starting from the notes on the left of the canvas, and including any pauses. If only partial payback is desired, the user selects part of the strokes (notes), then touches the replay icon, and only the selected notes are played back (partial replay). Furthermore, by turning the "rabbit and tortoise" dial on the palette, the tempo of the replay can be changed.

Implementation

The hardware of the *Vuzik* system consists of a canvas, brush and palette (Figure 2). For the canvas, a touch enabled wall-sized interactive display, we assembled the following product setups: front projection interactive whiteboard (94" SMART Board SB690) + projector (EPSON EB-1730W). For the brush we used unaltered, commercially available painting brushes, as the panels are pressure sensitive. For the palette, we cut a commercially available wooden painter's palette to a size that fits a child's hand and embedded two touch sensors (Phidget Circular Touch Sensor 1016), two rotation sensors (Phidgets Rotation Sensor 1109), and a slider sensor (Phidgets Slider Sensor 1112).

The software broadly comprises two packages, a WPF based application implemented in Microsoft Visual Studio 2008 C#, and a program implemented using Cycling '74 Max 5.1.7. Max is a graphical integrated development environment for sound and images, and is software that is widely used by musicians and new media artists.

EVALUATION

We examined the effect that larger gestural interaction has on children's creative activity by carrying out a comparative experiment using existing systems.

Study Design

Fostering an internal understanding: does the system encourage forming an overall idea of music in children?

As mentioned in the introduction, the formation of an overall conceptual understanding of one's creation is important in creative activity. Most traditional design support tools can be considered to inhibit the grasp of an overall idea of their creation due to the user becoming

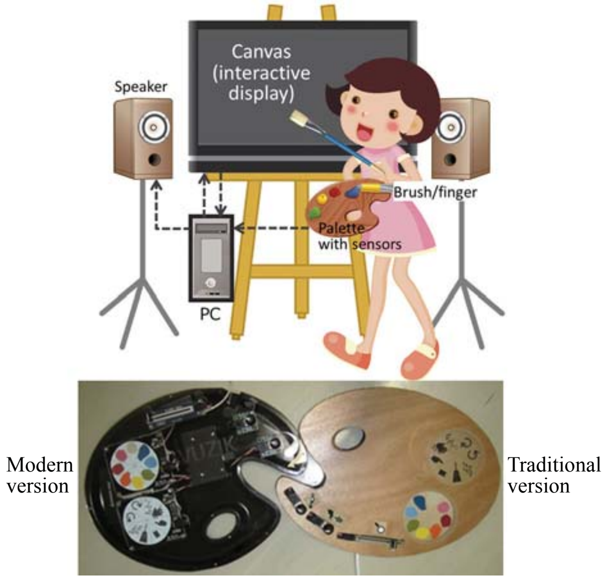


Figure 2. *Vuzik*'s hardware configuration and palette.

easily mired in detailed fine-adjustments. Our experiments set out to evaluate whether the user can express and construct music while forming an overall concept or idea. We will measure the system's effects on both the music generated by the user (the output) and the process in which the user generates the music (the process).

First, we investigate a method to evaluate whether the generated music (the output) has been formed with an overall idea in mind. Music is not mere sound, a simple melody or rhythm, but has a structure that expresses thoughts or emotions [32]. Western music, which is at the basis classical music in the world today, is temporally structured in a continuous dimension and a simultaneous dimension. A representative feature characterizing the structure in a continuous plane is phrasing [1], and a representative feature characterizing the structure in a simultaneous plane is harmony [2] [32]. We introduce these two, phrasing and harmony, as evaluation criteria. Three evaluators—students majoring in music—independently evaluated the pieces of music that each participant generated in task D (see below). They rated them on a three-point scale (2 points, 1 point, 0 points), and gave the participants a score averaging the ratings.

Then, to evaluate whether an overall idea was kept in mind during the process of creating the music, we introduce the macro operation ratio, or the extent to which the children operated the application in a dynamic manner without getting into the detailed fine-adjustment operations. Looking at three operations, namely the moving, transforming, and replay of tone objects, we classified the operations into macro-operations and micro-operations and calculated the ratio. For the operation classification criterion, we set the threshold

[1] Dividing the melody in natural groups made up of a number of notes.

[2] Structuring music using its parts such as melody, chords, rhythm.

relating to moving and transformation operations according to whether the rectangle surrounding the object after the user's operation would fit inside an area covered by double the size of the rectangle before the user's operation. If it did fit, we designated it as a micro-operation. If not, we designated it as a macro-operation. For the replay operation, we defined replaying the complete piece as a macro-operation, and replaying a specific note object that was selected more than once, a micro-operation. Moreover, in order to assess whether detailed adjustments were encouraged or not, we asked participants for their subjective evaluation. We used a 5-point Likert scale for the subjective evaluations.

Ease of learning: is the system easily mastered by children?
As described in the beginning, Dourish states that "because, in a digital environment where embodiment has been given consideration, the environment interacts in collaboration with one's own body movement, people can experience it through their senses and actions. [7]" It is also accepted that, in general, people find it easy to retain a memory of things they have actually experienced through their senses or actions. As a measure to evaluate the ease of learning for the application in this experiment, we use the frequency with which participants asked the experimenter questions about its operation during the experiment.

User experience: what do children think of their experience?
Finally, we need to confirm how the user's state of mind, cognition and emotions are affected as a result of using the system. In this study, we asked participants for their subjective evaluation in relation to enjoyment and satisfaction. We used a 5-point Likert scale for the subjective evaluation.

Furthermore, in order to objectively evaluate the children's emotions, we captured their facial expressions during the experiment. However, as *Vuzik* makes use of a projector, the environment with minimum lights must be maintained, and therefore, the recorded images containing numerous data pertaining to the children's expressions are difficult to distinguish. As a result, we were unable to give emotion ratings based on expressions using a tool such as the Smileyometer [24].

Experimental Conditions
We set up conditions for two tools, namely *Vuzik* and Hyperscore (Figure 3). We set the tool conditions to be between-subjects factors. Even though the number of participants was low at fourteen, we decided on between-subject factors due to consideration of the burden on the children who formed the participant group. Through an earlier informal user study it became clear that children would need a certain amount of time to arrive at the stage where they used a tool with which to manipulate sound through drawing, such as *Vuzik* or Hyperscore, as a tool not for mere drawing but to express music. Therefore we thought that, with the participants being children who can't maintain concentration as long as adults, asking one participant to use two tools for a long time may constitute a significant burden.

Our reasons for selecting Hyperscore as a control for comparison with *Vuzik* are as follows. First, Hyperscore is a tool that, similar to *Vuzik*, supports creation and arrangement of music by children through the freehand drawing of lines. Its functionality also has many similarities with *Vuzik*. Furthermore, Hyperscore was developed based on a theoretical study, and that is, generally, widely-used in many countries as music education software.

While *Vuzik* and Hyperscore are functionally similar, differences can be seen in the input/output interface. Comparing both systems, we will outline the common features as well as the differences from the perspective of a MVC (Model-View-Controller) model [4], used in designing user interfaces with the Smalltalk-80 programming language. From the Results section onwards we will examine the experiment results based on the three differences shown in the Table 1, i.e., diff1, diff2, and diff3.

Experimental Environments
Under the *Vuzik* conditions, participants stood in front of the canvas and expressed music by drawing lines on the

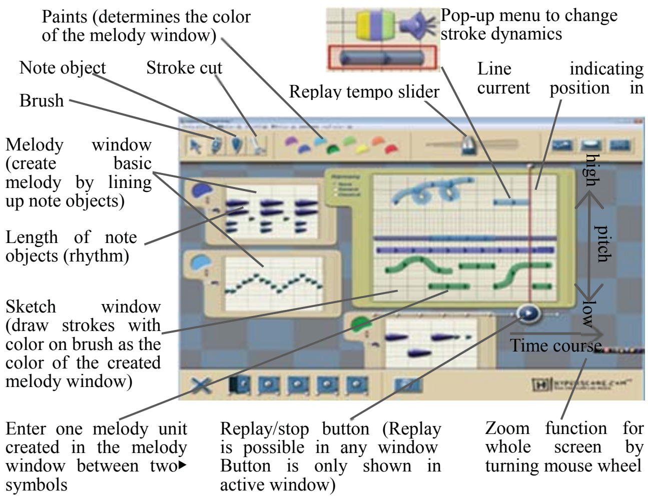


Figure 3. Hyperscore screen and functions.

	Common features	Differences
Model	● Data (basic elements of music: rhythm, pitch, dynamics, timbre) ● Processing (setting rhythm, pitch, dynamics, timbre, and replay tempo, replay/stop, etc)	(no big differences)
View	● Method of representing music (horizontal axis represents time, rhythm is expressed in lines or points, the vertical axis represents pitch, timbre is expressed in color, etc)	● Physical display size（diff1） *Vuzik*：94" Hyperscore：24" ● Window configuration（diff2） *Vuzik*：single window Hyperscore：multiple windows
Controller	● Operation by directly touching the Controller	● Number of Controllers（diff3） *Vuzik*：Each type of processing through corresponding individual controllers (button or brush) Hyperscore ： All processing operated through single Controller (mouse)

Table 1. Comparing *Vuzik* and Hyperscore.

Vuzik Hyperscore

Figure 4. Experimental setting.

Vuzik Hyperscore

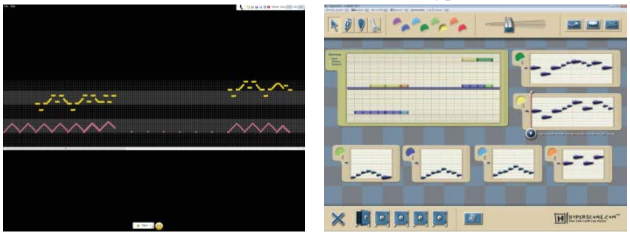


Figure 5. Task B (Fill-in-the-middle section task).

Vuzik Hyperscore

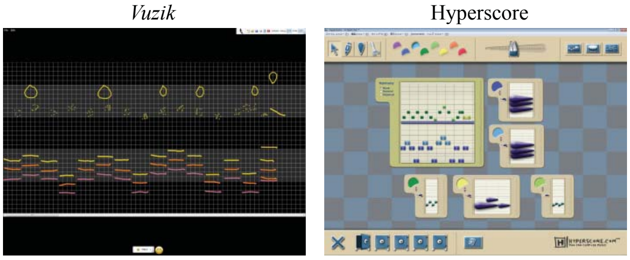


Figure 6. Task C (Supply-the-melody task).

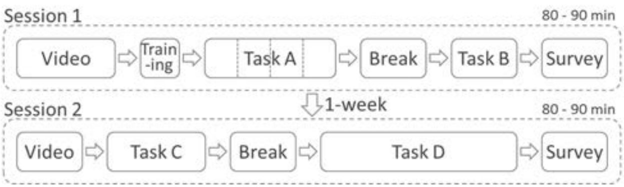


Figure 7. Experimental procedure.

canvas using a finger or a brush with the palette (Figure 4, left). Hyperscore music by drawing lines on the canvas using a mouse or keyboard (Figure 4, right). During the task, the display screen was recorded and the participants' facial expressions and all their movements were captured from two viewpoints.

Participants

Fourteen (5 female) elementary school children (4th-6th grade) from a public school in Japan recruited from the general public took part in the experiment. Recruitment was done through posters displayed within the school grounds and flyers distributed through the school's teachers. Participants were recruited if they fulfilled the conditions of liking to sing, perform and listen to music.

Tasks and Experimental materials

We prepared a total of four tasks, A, B, C and D. All music in tasks A, B and C was composed by one of the authors, who is a musician. We gave careful consideration to the music prepared, so that both *Vuzik* and Hyperscore had nearly identical music (in terms of melody, tempo, and rhythm).

Task A (Arrangement task): Participants were to change a provided piece of music. They were instructed to make at least two changes of their choice per piece. They could change the rhythm, pitch, dynamics and timbre of any note.

Task B (Fill-in-the-middle section task): Participants were to create the middle section for a piece in three parts that provided the first and last part (Figure 5).

Task C (Supply-the-melody task): Participants supply a melody for a piece that only contains a bass part in the low tone range and an accompaniment part in the high tone range (Figure 6).

Task D (Composition task): Task D is a completely open challenge. Participants were instructed to freely create a piece of music from scratch.

Experimental Procedure

We carried out the experiment following the procedure shown in Figure 7. Each participant took part in two sessions, which took place a week apart, in either one of the software tool setups.

For Session one, participants first viewed an approximately five minute long video explaining the software tool's operation. We then asked them to use the tool freely as a form of training in order to familiarize themselves with the tool. After that, we asked them to complete Task A. Task A contained four pieces of music. We limited the time per piece to four minutes, but we told the participants that they were allowed to finish before the time was up (same for tasks B, C, D). After a ten-minute break we asked them to complete Task B. Task B consists of one provided piece of music, to be completed within a time limit of ten minutes. After that, we asked them to answer a 5-point Likert scale questionnaire. The overall time required for Session One was approximately 80 to 90 minutes.

For Session two, participants first viewed an approximately ten-minute long video with a simple explanation of the basics of musical composition. We then asked participants to complete Task C. Task C consists of one piece of incomplete music to complete within a time limit to 15 minutes. After a ten-minute break, we asked them to complete Task D. Task D consists of the creation from scratch of one piece of music, with a completion time limit of 30 minutes. After performing the task, we asked participants to answer the exact same questionnaire as in Session One. The overall time required to perform Session Two was approximately 80 to 90 minutes.

We also told the participants that they could ask questions if they didn't understand something, or take a break if they became tired during the task.

RESULTS AND CONSIDERATIONS

Forming an Overall Mental Idea of the Music

First, we will report the results as to whether the generated music (output) was constructed with an overall mental idea of the desired creation in mind or not. Figure 8(a) shows the average scores and standard errors for phrasing and harmony – the indices to measure musical

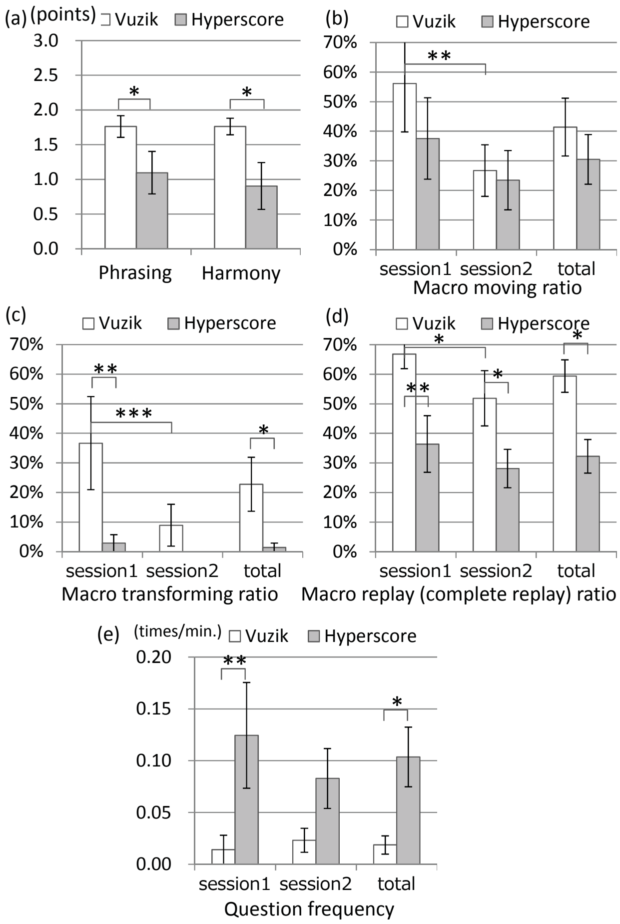


Figure 8. Quantitative evaluation results.

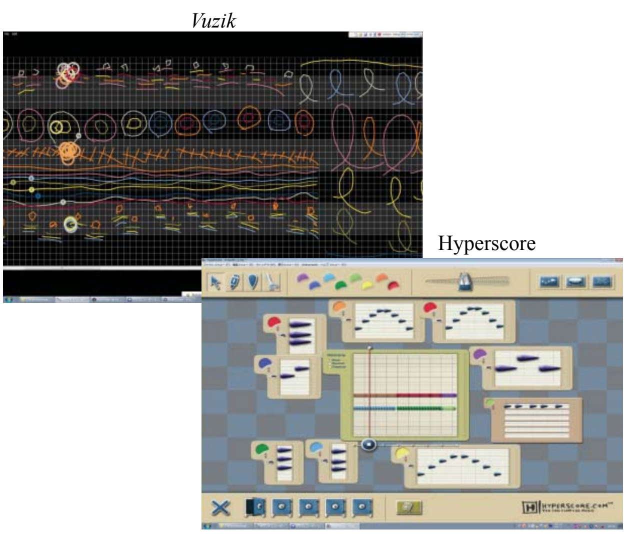


Figure 9. Example of pieces completed by participants in Task D.

structure – in the pieces participants completed in Task D. Figure 9 shows examples of pieces completed by participants in Task D. We carried out a significance test with a Welch's t-test with a significance level of 95%. *Vuzik* participants had significantly higher scores than Hyperscore participants for both phrasing (*Vuzik* (hereafter, "V"): *M*=1.76, *SD*=0.39; Hyperscore (hereafter, "H"): *M*=1.10, *SD*=0.75; t(8.98)=1.935, p<=.05) and harmony (V: *M*=1.76, *SD*=0.29; H: *M*=0.91, *SD*=0.83; t(7.48)=2.38, p<.05).

Next, we show the results as to whether the participant was keeping an overall idea of the music they were creating in mind in the process of creating music. Figure 8(b) shows the average values for the macro operation ratio (%). We obtained the macro operation ratio for the three operations of moving, transforming and replay of tone objects. We applied an arcsine transform on the percentage data and performed a tool (2: *Vuzik*, Hyperscore) X session (2: Session 1, Session 2) two-way ANOVA, in which session was repeated and tool was treated as a between-subjects factor.

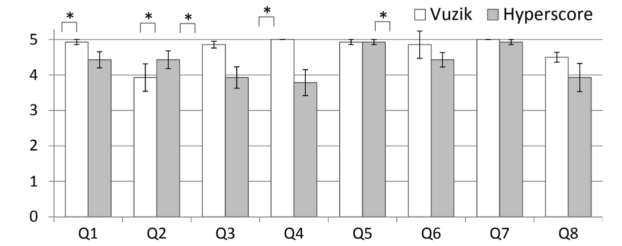
There was a significant main effect of session with respect to the moving operation (Session 1: 42.4 vs. Session 2: 26.3, F(1,12)=5.521, p<.05). There were no other significant effects or interactions. As the main effect of session was significant, we performed a post-hoc test (Tukey's HSD), which showed that in the *Vuzik* group the ratio of macro moving operations was significantly lower in Session 2 than in Session 1 (p<.01). Studying the data in detail, it was clear that the ratio was particularly low in the second half of Task D, which was a lengthy musical composition task.

For the transforming operation there were a significant main effect of tool (V: 22.0 vs. H: 1.9, F(1,12)=4.909, p<.05) and session (Session1: 19.2 vs. Session2: 4.7, F(1,12)=6.753, p<.05) (Figure 8, (c)). We saw no interaction between tool and session. Because the main effect of tool was significant we performed post-hoc test, which showed that the macro operation ratio was significantly higher in the *Vuzik* group than in the Hyperscore group (p<.05). As the main effect of session was also significant, we performed post-hoc test which showed that, similar to the moving operation, in the *Vuzik* group the macro operation ratio was significantly reduced in Session 2 (p<.001).

For the replay operation, there was also a significant main effect of tool (V: 49.6 vs. H: 32.9, F(1,12)= 5.290, p<.05) (Figure 8, (d)). There were no other significant effects or interactions. As the main effect of tool was significant we performed post-hoc test, which showed that the macro operation ratio was significantly higher in *Vuzik* than in Hyperscore (p<.05).

Figure 10 shows the results of the subjective evaluation questionnaire (Q1-Q4) as to whether fine adjustment operations were encouraged by the tools. We performed a tool X session two-way ANOVA in which session was repeated and tool was treated as a between-subjects factor. The graph represents the averages of the scale derived from participant's answer for two sessions. The results show that there was a main effect of tool in three out of the four questions, and that the *Vuzik* participants felt that they were able to boldly perform tasks without worrying about details more than the Hyperscore participants (Q1: V: 4.9 vs. H: 4.4, F(1,12)=7.350, p<.05; Q2: V: 3.9 vs. H: 4.4, F(1,12)=0.845, p=.0376; Q3: V: 4.9 vs. H: 3.9, F(1,12)= 4.783, p<.05; Q4: V: 5.0 vs. H: 3.8, F(1,12)=5.192, p<.05). We found no significant main effect of session or interaction between tool and session.

The aforementioned results suggest that on both aspects of the music generated (output) and the creation process,



Question as to whether you operate in a bold manner, without becoming caught up in detailed adjustments?	Q1	I drew lines resolutely, not nervously.
	Q2	I often thought that I'd like to move a line a change a shape a tiny bit.
	Q3	I could freely draw lines without caring about the details.
	Q4	I could freely make music without caring about the details.
Pleasure	Q5	I enjoyed making music.
	Q6	I enjoyed drawing lines.
	Q7	I enjoyed today's experiment.
Satisfaction	Q8	I was satisfied with the music I made.

Figure 10. Results of subjective questionnaire.

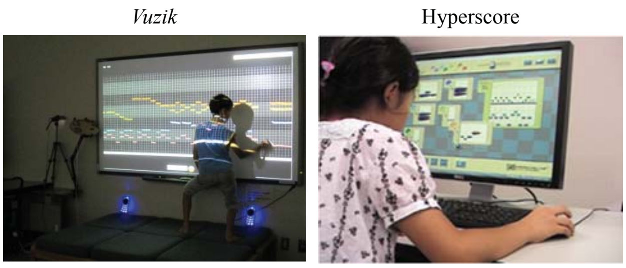


Figure 11. Participants during the experiment.

Vuzik encouraged the formation of an overall mental idea of their music while participants engaged with the expression and construction of their music, to a greater extent than with Hyperscore. We can say that, in spite of the feature that Hyperscore provides a window configuration which seems to facilitate the user's awareness of the musical structure (see Figure 3), the fact that results show *Vuzik* as more effective is a significant finding. However, dynamic operations decreased significantly in Session 2 containing the lengthy Task D with children in the *Vuzik* group, too. This indicates that with the lapse of time in creative activities the degree to which children operate with boldness decreases. This may indicate that *Vuzik* is particularly effective in the initial stages of creative activity.

Furthermore, in our observations regarding the video material, we frequently saw *Vuzik* users drawing lines with both feet spread wide apart while stepping from left to right, and *Vuzik* users who consciously or unconsciously put a mental and physical distance between their work and themselves by stepping away from the display or sitting down (Figure 11, left). On the other hand, in Hyperscore users, we observed a tendency to stay in a fixed position, and stare at the screen with a hunched posture (Figure 11, right). When moving strokes or when drawing users were also seen, with both tools, to grab hold of an object and hesitate for a moment before starting an action. However, once they started the action their appearance changed. We observed *Vuzik* users often moving objects or drawing resolutely while moving their arm or their entire body. On the other hand, many Hyperscore users carefully moved objects and looked as if they were continuously hesitating. This suggests that, when comparing *Vuzik* and Hyperscore, the main factor in encouraging the participants to form an overall conceptual understanding of their music is derived from the differences between the two (Table 1), by the difference in physical display size (diff1).

Ease of Learning

Figure 8(e) shows the frequency at which participants asked questions about the operation during the two sessions. We obtained for each session the number of questions per minute and performed a tool X session two-way ANOVA. There was a significant main effect of tool (V: 0.02 vs. H: 0.10, $F(1,12)=5.452$, $p<.05$). We found no significant main effect of session or interaction between tool and session. Since the main effect of tool was significant, we performed a post-hoc test which showed that the question frequency for *Vuzik* users was significantly lower than for Hyperscore users ($p<.05$). Based on these results we confirmed that, at least in regards to the frequency of questions relating to operation, ease of learning is higher with *Vuzik* than with Hyperscore. This is also consistent with the results of Q9 (see below) where *Vuzik* users felt that lower mental demand was required compared with Hyperscore users.

We discuss here the main factors resulting in *Vuzik*'s higher degree of ease of learning compared to Hyperscore on the basis of the differences between the two as shown in Table1. When we examined the content of the questions from participants in the Hyperscore group, most questions were related to the relationship between windows, followed by those related to which icon to select in order to perform heir desired task. This suggests that the difference is caused by the difference in window configuration (diff2) and the difference in the number of Controllers (diff3). Regarding diff2, there is the possibility that *Vuzik*'s simple window configuration reduces the cognitive burden in relation to system operation. As for diff3, we infer that the different physical buttons (Controllers) for each function in *Vuzik* encourage a different experience of embodiment, and whether as a result it was more easily retained in memory. On the other hand, we were unable to determine whether it was caused by the difference in physical display size (diff1) from the question content obtained.

User Experience

Figure 10 shows the results of the subjective questionnaire relating to pleasure, satisfaction and workload. We performed a tool X session two-way ANOVA.

For the questions relating to pleasure (Q5-Q7), there was a main effect of tool (Q6: V: 4.9 vs. H: 4.4, $F(1,12)=4.909$, $p<.05$) only for Q6 ("it was fun drawing lines"). We performed post-hoc test for Q6, and results showed that values were significantly higher for the *Vuzik* group ($p<.05$). We found no significant main effect of session or interaction between tool and session. For either Q5 or Q7, there was no significant main effect of tool. It is clear from the results of Q5 and Q7 that participants in both *Vuzik* and Hyperscore greatly enjoyed making music.

Regarding satisfaction (Q8), there was no significant main effect for either tool or session, even though the average values for participants in the *Vuzik* group were higher than that in the Hyperscore group. These results suggest that both *Vuzik* and Hyperscore participants were satisfied with the music they made.

In summary, there was little difference in the experience that users derived from *Vuzik* or Hyperscore. Regarding the positive effects of pleasure and satisfaction, *Vuzik* and Hyperscore users enjoyed the music expression and composing activity and were satisfied with their results to almost the same extent.

Summary

We evaluated *Vuzik*, an interface that allows children to interact with a large display using a brush or a finger, in comparison to Hyperscore, where people interact with a small display using a mouse. Our evaluation had three main perspectives: overall mental idea formation, ease of learning, and user experience.

When comparing *Vuzik* and Hyperscore in regards to the first aspect of overall idea formation, we found that *Vuzik* promotes overall conceptual understanding in children's through creative activity. The music that was created with *Vuzik* showed evidence that it was created with a mental image of the music as a whole (i.e., the melody or a part) in mind. Furthermore, when looking at the process, results showed that the proportion of bold operations was significant for *Vuzik* users. In other words, the proportion of detailed operations is more significant for Hyperscore users. Based on observations of video recordings of the study, these results could be because *Vuzik* users kept the overall mental image of their music in mind while, through large gesture interaction, they drew in a dynamic manner and put distance between themselves and their work during the creation process. This finding raises the possibility that the use of large-size displays, which inevitably encourages larger-scale bodily interaction, is effective not only in our *Vuzik* system, but in digital environments supporting children's creative activities, in general. A comparative experiment with *Vuzik* displays of varying sizes is required to closely verify the effectiveness of large-scale gestural interaction, and is left to be pursued in future work.

On the other hand, for *Vuzik* users, the proportion of bold operations decreased greatly during the second session of our study, which included a lengthy task, compared with the first session. While this shows that large gesture interaction as used in *Vuzik* is effective in the first, more explorative stage of the creative activity, at the same time it indicates the possibility that it may not be the optimum form of interaction in the intermediate and final stages of creative activity, which may require more meticulous attention to details. Creative activity consists of a number of stages, and exploring system requirements that can provide the appropriate interface for each stage will be one of the major design challenges we will examine in the future.

Reflecting on the second aspect of ease of learning, we confirmed that, compared to Hyperscore, the frequency of operation-related questions was significantly lower for *Vuzik*. A qualitative analysis of the question content suggests that a simple window configuration reduces cognitive burden on system operation, and that different experiences of embodiment, namely operation of different physical buttons for each function, promote recognition and thus lessen the need to memorize operations.

In regards to user experience, no major differences were observed between the *Vuzik* and Hyperscore, and we confirmed that users of both systems thoroughly enjoyed the music expression and creation activities and were satisfied with their experiences and composition outcomes to almost the same extent.

CONCLUSION

In this paper we presented *Vuzik*, an interactive system aimed at children, supporting creative musical activity by the use of tangibles and larger scale gestures in front of a large interactive display. The results of a comparative study between *Vuzik* and Hyperscore, a more traditional WIMP interface where children compose music using a desktop PC, showed that *Vuzik* is effective in encouraging an overall conceptual understand of their creation, considered to be important in creative activity. This suggests that, when supporting of children's creative activity, smaller interactive spaces and lack of rich physical mappings and engagement can be significant shortcoming. Our findings confirm previous research on the effectiveness of embodied interaction in a large interactive display space when supporting creative activity.

To substantively improve *Vuzik*, a deeper understanding of the effect that large-scale gestural interaction has on creative activity is necessary. This paper focused on overall idea formation of one's creation as an important aspect in creative activity, but does not cover the entirety of creative activity. In the area of creative activity the importance of metacognitive activity has been pointed out, where one reflects upon one's own actions (e.g. Dewey [6] and Schön [28]). Systems which support such reflective thinking activity are also being explored in the HCI field (e.g. [19]). We see great potential in future work on including adding reflective thinking elements in creative interfaces such as *Vuzik*.

Another future work aspect we are planning to explore is investigating *Vuzik*'s potential as a tool for collaborative music expression and creation by multiple users. The expression of music is not only essential for the individual, but is also an important communication activity. When we asked several children to informally use *Vuzik* together, we observed a chain of events where their imaginations were stimulated by the output from other children interacting with the system, which in turn became a source of inspiration and gave rise to further new output.

Another, more practical layer of our vision for *Vuzik* is developing classroom kits which will allow teachers in classrooms already equipped with large interactive displays (not uncommon in the developed world public educational system) to use ready-to-use in-class experiences and activities which explicitly promotes children's musical creativity skills.

ACKNOWLEDGMENTS

Part of this study was funded in part by a JSPS Grant-in-Aid for Scientific Research (24600003), NSERC, Surfnet, GRAND, AITF, and SMART Technologies. We also want to express our gratitude for the cooperation of children of Chofu Daiichi Elementary School of Japan.

REFERENCES

1. Benveniste, S., Jouvelot, P., and Péquignot, R. The MINWii Project: Renarcissization of Patients Suffering from Alzheimer's Disease Through Video Game-Based Music Therapy. *ICEC 2010*, LNCS 6243 (2010), 79–90.
2. Black, A. Visible Planning on paper and on screen: The impact of working medium on decision-making by novice graphic designers, *Behaviour and Information Technology*, 9, 4 (1990), 283-296.
3. Bott, J.N., Crowley, J.G., and LaViola, J.J. Exploring 3D gestural interfaces for music creation in video games, *Proc. FDG 2009* (2009), 18-25.
4. Burbeck, S. Applications Programming in Smalltalk-80(TM): *How to use Model-View-Controller (MVC)*, (1992).
5. Cao, X., Lindley, S.E., Helmes, J. and Sellen, A. Telling the whole story: Anticipation, inspiration and reputation in a field deployment of TellTable, *Proc. CSCW2010*, ACM Press (2010), 251-260.
6. Dewey, J. *How We Think*, D.C.Heath, 1910.
7. Dourish, P. Where the Action Is: The Foundations of Embodied Interaction, MIT Press (2001).
8. England, D., Randles, M., et al. Towards an Advanced Framework for Whole Body Interaction. *Virtual and Mixed Reality*, 5622 (2009), 32-40.
9. Farbood, M., Pazstor, E., Jennings, K. H: A graphical Approach to Composing Music, *Proc. IEEE CG&A*, Special Issue on Emerging Technologies (2004), 50-54.
10. FlexiMusic: FlexiMusic Kids Composer, http://fleximusic.com/product/fleximusic-kids-composer.
11. Halpern, M., Tholander, J., et al. MoBoogie: Creative Expression Through Whole Body Musical Interaction, *Proc. CHI 2011*, ACM Press (2011), 557-560.
12. Ichino, J., Pon, A., Sharlin, E., Eagle, D. and Carpendale, S. Vuzik: Creative Music Expression for Children thorough Whole Body Interaction, *IPSJ Journal*, 53, 12, 2773–2786, 2012.
13. intel: Dragencer, http://www.intel.com/.
14. Knörig, A., Müller, B., Wettach, R. Articulated Paint: Musical Expression for Non-Musicians, *Proc. NIME 2007*, ACM Press (2007), 384-385.
15. Merrill, D., Raffle, H. and Aimi, R. The Sound of Touch: Physical Manipulation of Digital Sound, *Proc. of CHI 2008*, ACM Press (2008), 739-742.
16. McCarthy, J. and Wright, P. *Technology as Experience*, 2007.
17. Montemayor, J., et al. Tools for children to create physical interactive storyrooms, *Computers in Entertainment (CIE)*, 2, 1, ACM Press (2004).
18. Nakakoji, K. and Yamamoto, Y. Interaction Design Sketches - A Case Study in the ARTware Project -, *Proc. Design Symosium 2010*, 2010.
19. Nakakoji, K., et al. Toward Unweaving Streams of Thought for Reflection in Early Stages of Software Design, *IEEE Software*, 29, 1 (2012), 34-38.
20. Paradiso, J. A., Hsiao, K.-Y., Strickon, J. and Rice, P. New Sensor and Music Systems for Large Interactive Surfaces, *Proc. ICMC 2000* (2010), 277-280.
21. Petersen, M.G., Iversen, O.S., et al. Aesthetic interaction, *Proc. DIS'04*, ACM (2004), 269-276.
22. Pon, A., Ichino, J., Eagle, D., Sharlin, E., D'Alessandro, N. and Carpendale, S. *Vuzik*: A Painting Graphic Score Interface for Composing and Control of Sound Generation, *Proc. ICMC 2012* (2012), 579-583.
23. Raffle, H., Vaucelle, C., Wang, R. and Ishii, H. Jabberstamp: Embedding sound and voice in traditional drawings, *Proc. IDC 2007*, ACM Press (2007), 137-144.
24. Read, J. C., MacFarlane, S. J., & Casey, C. (2002). Endurability, engagement and expectations: measuring children's fun, *Proc. IDC 2002*, ACM Press (2002), 189-198.
25. Rosenbaum, E., et al. Singing Fingers: fingerpaint with sound, *Proc. IDC 20010*, ACM (2010), 308-310.
26. Ryokai, K., Marti, S., and Ishii, H. I/O brush: drawing with everyday objects as ink, *Proc. CHI 2004*, ACM Press (2004), 303-310.
27. Ryokai, K. and Cassell, J. StoryMat: A Play Space for Collaborative Storytelling. *CHI Ext. Abstracts '99*, ACM Press (1999).
28. Schön, D. A. The Reflective Practitioner: How Professionals Think in Action, Basic Books, 1984.
29. Strachan, S., et al. BodySpace: Inferring body pose for natural control of a music player, *CHI Ext. Abstracts 2007*, ACM Press (2007), 2001-2006.
30. Subotnick, M. MakingMusic, http://www.creatingmusic.com/.
31. Tano, S. Analysis of Obstruction and Promotion of Human Creative Work by Information Systems, *Proc. Human Interface Symposium '99*, Human Interface Society (1999), 791-796.
32. Umemoto, T. *Children and Music*, University of Tokyo Press, 1999.
33. Volpe, G., Camurri, A. A system for embodied social active listening to sound and music content, *ACM Journal on Computing and Cultural Heritage (JOCCH)*, 4, 1, Article 2, ACM Press, 2011.
34. Williams, A., et al. From Interaction to Participation: Configuring Space through Embodied Interaction, *Proc. Ubicomp 2005* (2005), 287-304.
35. Zigelbaum, J., Millner, A., Desai, B. and Ishii, H. BodyBeats: whole-body, musical interfaces for children, *CHI Ext. Abstracts 2006*, ACM Press (2006), 1595-1600.