Multimodal Design for Enactive Toys

A. de Götzen, L. Mion, F. Avanzini[1], and S. Serafin[2]

[1] Department of Information Engineering, University of Padova
http://smc.dei.unipd.it/
[2] Medialogy Department, Aalborg University in Copenhagen, Denmark
http://www.imi.aau.dk/
{degotzen,luca.mion,avanzini}@dei.unipd.it, sts@imi.aau.dk

Abstract. In this paper we will investigate how non–visual senses can be used in toys to enhance and enrich the play experience of all children, while favoring accessibility and inclusion of visually-impaired children. Previous research has shown that – especially for young children developing sensory-motor skills – exploration and play are two tightly linked activities: everything is new and needs to be "investigated" and playful behaviors emerge from active exploration. We will propose a new approach in designing and creating objects that elicit this type of behavior and encourage exploration by providing real–time dynamic, haptic, tactile, auditory, and even olfactory feedback depending on children's gestures, movements, and emitted sounds. We believe that this design paradigm is highly innovative with respect to previous research and existing products – whose interaction is very often based on static feedback. Interactive and dynamic feedback is intrinsically more engaging and allows a variety of quality learning patterns.

1 Introduction

According to the traditional mainstream view, perception is a process in the brain where the perceptual system constructs an internal representation of the world, and eventually action follows as a subordinate function. Two assumptions emerge from this view. First, the causal flow between perception and action is primarily one-way: perception is input from world to mind, action is output from mind to world, and thought (cognition) is the mediating process. Second, perception and action are merely instrumentally related to each other, so that each is a tool for the other. Recent theories have questioned such a modular decomposition and have rejected both the above assumptions: the main claim of these theories is that it is not possible to disassociate perception and action schematically, and that every kind of perception is intrinsically active and thoughtful. As stated in [1], only a creature with certain kinds of bodily skills (e.g. a basic familiarity with the sensory effects of eye or hand movements, etc.) can be a perceiver. One influential contribution in this direction is [2]. The authors present an "enactive conception" of experience, which does not occur inside the animal, but is rather something that the animal *enacts* as it explores its environment. In this view, the subject of mental states is the *embodied* animal, situated in

R. Kronland-Martinet, S. Ystad, and K. Jensen (Eds.): CMMR 2007, LNCS 4969, pp. 212–222, 2008.

the environment. The animal and the environment form a pair in which the two parts are coupled and determine each other. The term "embodied" highlights two points: first, cognition depends upon the kinds of experience that are generated from specific sensorimotor capacities. Second, these individual sensorimotor capacities are themselves embedded in a biological, psychological, and cultural context. The enactive knowledge is then stored in the form of motor responses and acquired by the act of "doing" [3].

Examples of enactive knowledge are represented by the competences required by tasks such as typing, playing a musical instrument, sculpting objects, whistling, tying shoelaces etc. This type of knowledge transmission can be considered natural and intuitive, since it is based on the experience and on the perceptual responses to motor acts and it involves more than just one modality of interaction. Multimodal interaction for children however poses new specific challenges. Conceivably, the kind of support that children need is different from that of adolescents and adults. Toys for children are very often poor in term of interactivity, while multimodal interaction should be the main way of exploring the environment and learning from it. The importance of sound as a powerful medium has been largely recognized, up to the point that there are objects on the market that reproduce prerecorded sounds by pushing buttons or touching areas. However, such triggered sounds are extremely unnatural, repetitive, and ultimately annoying. The same is often true for the tactile/haptic part of the interaction. Things may vibrate (and they have basic haptic properties due to the fact that they are made of physical materials) but designing a haptic interaction that can be dynamically changed still poses several challenges. As a consequence the interaction is unrealistic and un-engaging, and the learning patterns are very stereotyped. The key for a successful exploitation of sounds in toys interfaces is to have models that respond continuously to continuous gestures, just in the same way as rattles or other physical sounding objects do when they are manipulated by children, eliciting the enactive exploration of the world through multimodal interaction and helping them to discover and recognize many different sounding gestures, each characterized by specific movement, force, velocity etc.

A third sense, which has rarely been explored in multimodal interfaces so far, is the sense of smell. Simple "scratch-and-sniff" cards are available in some toys and games, but these cannot be changed in response to the input from a child. Vision, on the other hand, has been much investigated. This is the main reason why the majority of available toys are not accessible to visually impaired children that very often have to use specially designed interfaces that are more educational instruments than amusing toys. This also implies that these instruments exclude the collaboration of sighted children in the play activity of their non sighted friends. We will focus on the non-visual senses trying, as a important side effect, to bridge the gap between educational interfaces and amusing toys, specially designed interfaces for non sighted children and toys designed for the abilities of all children.

We think that children need toys that are designed and created to be explored through all the senses: they should vibrate, they should smell, they should react to our feelings and gestures and they should respond to our actions. They should tell us where we are, what they are and what we are doing not just by looking at them but by squeezing them, smelling them, shouting to them, caressing them, letting them fall down, bouncing them one against the other etc.

2 What Do We Learn? Musical Instruments as a Compelling Example

Gesture and sound seem naturally connected in a clear and obvious way: the image of instrument players who learn to use their body in order to produce sound is indeed widespread and compelling enough. Musical gesture can be simply thought as a gesture that produces sounds in a continuous feedback loop: this is a general definition that can be used in many interactive contexts besides the musical ones. In this respect it can be useful to design interactive sounding toys that do not have to be the exact replica of original instruments, while they have to help in acquiring some basic musical skills. Children get tired soon of traditional teaching methods e.g. for bowing instruments, since they have to spend many hours before the teacher is satisfied by the sounds and modulations produced. In this case arc bowing can be taught - along with aesthetics - at a more gestural level by means of multimodal interfaces. Controllers, for instance, can be tuned in order to make children have fun during their learning, adding a visual/haptic feedback that can engage the child in exercising/playing with the instrument. Moreover very young babies could start their musical training with basic toys, learning that is not enough to kick an object to produce sound, but that it might sound in a more pleasant way by just caressing it, or finding the needed force to squeeze it.

The idea is then to teach musical gestures through simple interactions mediated by the child's body. Each object/instrument has its own way to be played: children can learn the effects of their gestures through an enactive exploration of the object, learning what are the 'musical gestures' needed to produce the sound they want. Sound and gestures are indeed very important for children in prescholar age. In fact, cognitive sciences focus on how humans interact with their environment, searching the connection between perception and action to bridge the semantic gap that humans experience in their everyday life: as sound and music are linked to their physical energy, the content of auditory information has to be linked to meaningful actions that we can use to access the encoded high-level information. This relation is crucial for non-verbal communication in general; implicit messages (e.g. expressive content) are indeed the basis of the communication process in different social situations, especially for children whose language is based on sounds and gestures, organized by semantics and constructs only at a later stage. Those sounds and gestures can be very expressive and rich of emotional content, as music can be.

Humans use recognition and expression of affect to detect meaning [4] and communication by means of vocalization, facial expression and posture, while gestures express affect (emotion) and convey information more powerfully and efficiently than spoken language. Concerning the communication between children, tactile/auditory perceptions are the major actors for emotional response and affect: the sound-making gestures of infants are the earliest attempts for separating basic emotions [5], and earlier exposure to sound patterning has profound effects on perceptual and emotional development, while deprivation can lead to future weak development of linguistic and musical skills [6]. The understanding of the emotional response related to sensory experiences and object relationships is then a crucial issue, and a novel design paradigm for expressive toys can exploit this idea of embodied-expressive knowledge; moreover, expressive paradigms based on affective and sensorial adjectives can be used to provide expressive feedback to children according to their input. Children can then associate well known feelings and basic emotions to auditory and multimodal feedback, expressed by physical metaphors which can be directly mapped to higher emotional labels [7]. Applications in this direction can be imagined for teaching/educating to musical gestures rather than to the musical language itself. Gestural skills can be developed by means of interfaces for controlling in real time the expressive information by tactile interaction and controllers to map and to transform audio data, simultaneously promoting and stimulating the communication process.

3 How Do We Learn? Touch, Movement and Sound

In light of embodied perception theories, it is clear that developing "enactive interfaces" implies developing techniques for multimodal feedback and input, including sound, touch and gesture. Sound and touch are inherently tied to movement. Without movement there would be no sound, and the sounds that we perceive are influenced by the way our ears move within the world. Most of the information received by touch is also a result of movement, this being particularly true for proprioception and kinesthesia. This is well known for children who explore the objects around them by touching, moving themselves and the objects, hearing the results of their actions etc. Therefore the study of haptic and auditory feedback is particularly interesting in this context, since we are focusing on the dynamic properties of the interaction and on the learning process that is elicited by the action.

Sound and haptic feedback in interaction are related in a number of different ways. Actions produce sounds by direct, physical manipulation of physical objects. There is a physical energetic consistency between action and produced sounds: sounds can be produced as a result of instantaneous object manipulation (the sound starts after the end of action), or as a result of continuous object manipulation (the sound continues during the manipulation). Everyday sounds are used to infer information from the environment, to know what things are, where they are, and what happens. They can be used to inform the environment about our actions or intentions, in order to show what we are doing, where and when we are doing it. Studies on the interplay between touch and audition concerning object properties have mainly focused on contact properties such as hardness, stiffness and texture. For surface roughness and stiffness it has been shown [8], that touch dominates over audition, but both of them can improve the perception or even create illusions (see [9] for an example of auditory-haptic illusion).

In light of these perceptual studies, simultaneous audio-haptic rendering is a particularly interesting problem in the development of enactive multimodal interfaces. Recent literature has proposed *physically-based models* for sound synthesis, i.e. sound synthesis algorithms based on a physical description of sound generating mechanisms. Since the resulting computational structures respond to physical input parameters, they automatically incorporate complex responsive acoustic behaviors. A second quality of physically-based approaches concerns interactivity and ease in associating motion to sound control. As an example, the parameters needed to characterize collision sounds, e.g. relative velocity at collision, can be directly used to control a physically-based model, and the sound feedback responds consequently to gestures and actions in a natural way. Various approaches have been proposed in the literature for contact sound modeling. Modal synthesis [10] was proposed in [11] as a framework for describing the acoustic properties of objects; the modal representation is naturally linked to many *ecological* dimensions of the corresponding sounds: modal frequencies depend on *shape* and geometry of the object, *materials* determine the sound decay characteristics, and so on. Physically-based models for real-time synchronous haptic-sound rendering is an approach that will ensure synchronization and perceptual similarity between haptic and audio feedback. A significant amount of recent literature deals with this problem. In [12] the modal synthesis techniques described in [11] were applied to audio-haptic rendering. A related study was recently conducted in [13]: physically-based sound models were integrated into a multimodal rendering architecture, and the setup was used to run an experiment on the relative contributions of haptic and auditory information to bimodal judgments of contact stiffness.

4 Inspiring Related Research

Pioneering works in the field of innovative toys for children are due to Seymour Papert, who developed the Logo programming language (the first children toys with built-in computation), and to Mitchel Resnick, whose research group developed the "programmable brick" technology that inspired the LEGO MindStorms robotics kit and the PicoCricket artistic-invention kit [14,15]. Existing applications in this area can be categorized according to broad keywords that are commonly encountered in commercial product and research: Education [16], Physical programming [17], Interactive story telling [18], Collaboration [19,20]. These general trends offer a large variety of applications that deal with the cognitive level: creating dancing creatures, animated stories, video games, and interactive/collaborative painters focus on cognitive processes based on a bottom-up communication of meanings [21]. Moreover, multimodal information provided by these toys is typically based on iconic messages, resulting in poor interaction, e.g. based on triggering some kind of recorded sample. None of the applications reviewed here is really related to enactive concepts, despite learning by doing is arguably the most efficient and effective form of exchange between e.g. the teacher and the student.

In the following section we will describe the typical interaction that will be provided by enactoys with the help of a couple of scenario examples, underlying the learning process that such toys can elicit.

5 Scenarios

The first years of a baby life are a continuous discovery: a child starts to learn which reactions he produces in the world around him, how things sound, move, smell and how they can be used. Everything can become a toy and the boundary between a tool, a toy, and a simple gadget is never clear and determined. In particular pre-scholar babies spend hours playing with very simple objects that become whatever they wish, according to their shape or properties. A large pillow can become a spaceship, while an empty box can be knocked with the hands or with a spoon and it can become a drum or the door of a little house. Any object can stimulate the imagination of babies, and the general rule is that the simpler the object is, the bigger are the transformations it can perform in the baby's mind. Children, objects and environment can be considered as the three key elements in a play scenario, as described by Garvey [22]: "They (ndr. objects) provide a means by which a child can represent or express his feelings, concerns, or preoccupying interests. (...) Further, for the child an unfamiliar object tends to set up a chain of exploration, familiarization, and eventual understanding: an often-repeated sequence that will eventually lead to more mature conceptions of the properties (shape, texture, size) of the physical world". Following these ideas, one has to create simple toy–objects that, once explored, exhibit their multimodal properties while suggesting basic and complex reactions and interactions and improving the learning process and the motor skills of children.

Design Paradigm N.1: ***Children impose their own meaning and mental imagery over toys and things: there is no need to explicitly suggest specific roles, functions and/or images with toys. Rather, toys should react to children actions, providing a sensible mix between redundant and unexpected information.***

The interaction with the world requires the integration of information through different senses and we should keep all the senses in mind while designing toys for children: different abilities and capacities should be stimulated and enhanced from the same tool, allowing visually impaired children and unimpaired ones to play together on equal terms. All these aspects are particularly important in building toys for visually impaired children:

Design Paradigm N.2: ***toys must be designed for children abilities rather than be compensating for their disabilities.***

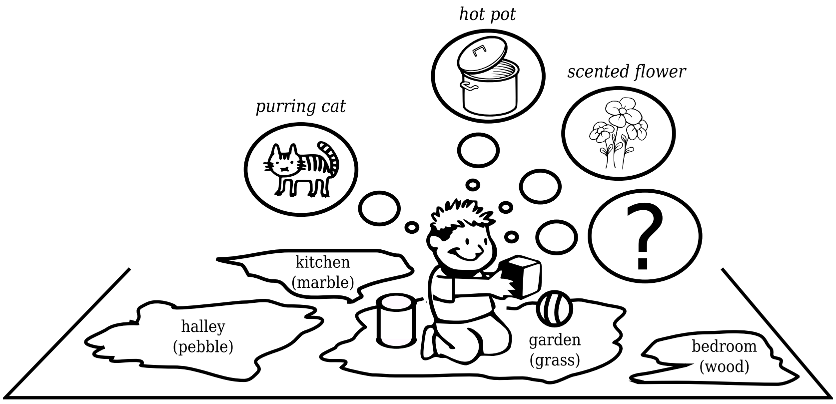

Fig. 1. The playground: simple objects enhanced by actuators and sensors can represent different environments

For all these reasons, we are interested in creating toys which present the following characteristics: (i) focused on enactive knowledge, (ii) enhanced by haptic, auditory and olfactory feedback to provide expressive output, (iii) embedded with sensors which trigger continuous interaction, (iv) reconfigurable to stimulate children's creativity and collaboration. All these characteristics are needed to provide a learn and play context that enhances the autonomy of the child and the collaboration with other children and that focuses on the abilities of the child.

5.1 Scenario 1: Playground

The playground is designed to stimulate and motivate the child to explore the environment in order to experience cause and effect of his/her actions and to learn from physical play and manipulation of objects their characteristics and their relationships. The aim is also to engage children through more full-body play patterns. The basic idea is to let the child explore a special carpet on top of which several simple objects (cubes, spheres) of different sizes and materials will be placed, as shown in Figure 1. The set-up sketched in this figure may represents, according to the initial settings, different environments: a kitchen, a garden, a beach, the bathroom etc. The child can discover what kind of environment she/he is exploring by walking, touching and shaking and smelling the objects, talking to them. Synthetically simulated footsteps may even be produced while the children wander in the different locations of the environment.

The set-up may be realized by using a kit similar to traditional Duplo Lego kits, enhanced by low cost sensing and control (e.g. phidgets technology[1]) in order to enhance the children's awareness of everyday life's multimodal feedback. According to the environment that it is simulated, the different objects may

[1] http://www.phidgets.com

Table 1. The kitchen playground scenario

Objects	Metaphors	Technology
carpet	parquet, broken glasses, liquid sounds or boiling water	pressure sensors
large cylinder	the blender, if you squeeze it, it will start blending	accelerometers, pressure sensors, RFID technology, synthetic receptors
small cylinders	spoon(wood) and knife(metal)	tactile actuators, equipment for measuring static or dynamic forces or torque
large cube	the pot (hitting the pot with the small cylinder will produce a percussive sound) moving the pot some water could fall	RFID technology, accelerometers, haptic and vibrotactile sensors
small cubes	dishes: if the child make them fall on the ground they will break	accelerometers, pressure sensors
big sphere	the dish-washing machine: the child can open and close it, put inside the smaller objects: shaking it he will identify the number of objects that are inside	RFID technology, vibrotactile sensors
small spheres	spices: if you shake them they smell like coffee or spices	smell actuators, controllers for measuring the humidity and the temperature

be programmed to assume different characteristics: in the Table 5.1 a "kitchen scenario" is described as example. Some of the objects may also assume animate and/or expressive behaviors: as an example, in a garden-like environment a simple cylinder will be a dog and will bark in different ways according to the child voice or gestures, interacting with him according to some expressive features extracted from the child's voice.

5.2 Scenario 2: The Reactoy–Band

Tangible user interfaces and more precisely, table based tangible interfaces in which digital information becomes graspable with the direct manipulation of simple objects available on a table surface, can fulfil many of the special needs required in designing inclusive toys. They allow an intimate and sensitive control, with a more macro-structural and higher level control which is intermittently shared, transferred and recovered between the children and the machine. Tabletop interfaces favor multi-parametric and shared control, exploration and multi-user collaboration, while they can also allow delicate and intimate interaction (e.g. moving and turning two objects with both hands). Seamless integration of visual and tactile feedback with physical control allows for natural and direct interaction. Designing for children abilities means to design toys to be fun and enjoyable also for non visually impaired children, encouraging in this way the

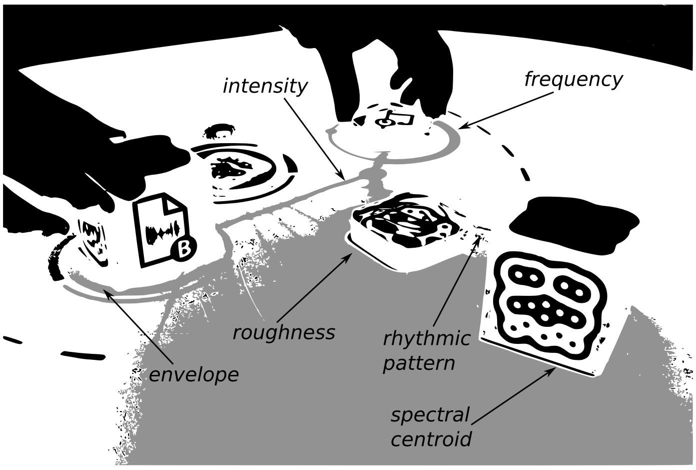

Fig. 2. Users interacting in the reactoy band scenario (the original picture of the reacTable device was taken from http://mtg.upf.es/reactable)

collaboration between children with different abilities as if they were different players in a common orchestra.

An example of a tabletop tangible musical interface is the reacTable [23]. Several musicians can share the control of the instrument by caressing, rotating and moving physical artifacts on a table, constructing different audio topologies in a kind of tangible modular synthesizer or graspable flow-controlled programming language. The objects interact with each other based on their proximity and their affinity, creating complex networks of flowing sound. Originally the resulting animated waveforms are projected from under the table, giving the performers a visual feedback which very reflects the sound flowing through the sonic network and all the inner workings of the sound generating mechanisms.

This kind of tangible musical interfaces can represent a perfect scenario for a collaborative toy that stimulates children to play a collaborative instrument that stimulates all their capabilities and collaboration by making music together around a table. The social affordances associated with tables directly encourage concepts such as "social interaction and collaboration" [24] or "ludic interaction" [25]. Tabletop interfaces should by definition include several simultaneous users, each using both hands and several objects at the same time. Visually impaired children may discover the function of each object perceiving its shape and the texture of its surface. They may understand the relationship between objects movements and effects on the music that they are hearing by simply manipulating the objects on the table. Vibro-tactile feedback, specific actuators and auditory feedback can give information about object states and table topology.

6 Conclusions

This exploratory work proposes a new way to think about toys for children, taking the particular perspective of multimodal interaction. We called these objects *enactoys* since they provide interaction based on the enactive paradigm, where multimodal feedback is intimately tied to action – i.e. the human is "in the loop". The enactoys can create new opportunities for playing and learning–through–play, as well as greatly improving accessibility and inclusion for children with special needs, who will be able to play on equal terms with sighted peers. Research in this direction has to combine state of–the–art technology on multimodal interfaces, techniques for the extraction of high-level expressive features from gestures and sound, and participatory design methodologies, in order to investigate how to design interactive objects that empower children to create their own environments and play patterns.

Acknowledgments. This research was partially supported by the EU Sixth Framework Programme – IST Information Society Technologies (Network of Excellence "Enactive Interfaces" IST-1-002114, http://www.enactivenetwork.org).

References

1. Noe, A.: Action in perception. MIT Press, Cambridge (2005)
2. Varela, F., Thompson, E., Rosch, E.: The Embodied Mind. MIT Press, Cambridge (1991)
3. Luciani, A., Florens, J., Castagne, N.: From action to sound: a challenging perspective for haptics. In: Workshop on Enactive Interfaces (Enactive 2005), Pisa, Italy (January 2005)
4. Picard, R.: Affective computing. MIT Press, Cambridge (1997)
5. Dore, J.: Feeling, form, and intention in the baby's transition to language. In: Golinkoff, R.M. (ed.) The transition from prelinguistic to linguistic communication (1983)
6. Sabbadini, A.: On sounds, children, identity and a 'quite unmusical' man. Br. J. Psychother. 14(2), 189–196 (1997)
7. De Poli, G., Avanzini, F., Rodà, A., Mion, L., D'Incà, G., Trestino, C., Pirrò, C., Luciani, A., Castagne, A.: Towards a multi-layer architecture for multi-modal rendering of expressive actions. In: Proceedings of 2nd International Conference on Enactive Interfaces (Enactive 2005), Genova, Italy (November 2005)
8. Lederman, S., Klatzky, R., Morgan, T., Hamilton, C.: Integrating multimodal information about surface texture via a probe: relative contribution of haptic and touch-produced sound sources. In: Symp. Haptic interfaces for virtual environment and Teleoperator Systems (HAPTICS 2002), Orlando, FL (2002)
9. Bresciani, J.P., Ernst, M.O., Drewing, K., Bouyer, G., Maury, V., Kheddar, A.: Feeling what you hear: auditory signals can modulate tactile tap perception. Exp. Brain Research 162(2), 172–180 (2005)
10. Adrien, J.M.: The missing link: Modal synthesis. In: De Poli, G., Piccialli, A., Roads, C. (eds.) Representations of Musical Signals, pp. 269–297. MIT Press, Cambridge (1991)

11. van den Doel, K., Pai, D.K.: The sounds of physical shapes. Presence: Teleoperators and Virtual Environment 7(4), 382–395 (1998)
12. DiFilippo, D., Pai, D.K.: The AHI: An audio and haptic interface for contact interactions. In: Proc. ACM Symp. on User Interface Software and Technology (UIST 2000), San Diego, CA (November 2000)
13. Avanzini, F., Crosato, P.: Integrating physically-based sound models in a multimodal rendering architecture. Comp. Anim. Virtual Worlds 17(3-4), 411–419 (2006)
14. Seymour, P.: The Children's Machine: Rethinking School in the Age of the Computer. Basic Books, New York (1993)
15. Wallich, P.: Mindstorms: not just a kid's toy. IEEE Spectrum 38(9), 52–57 (2001)
16. Hengeveld, B., Hummels, C., Voort, R., van Balkom, H., de Moor, J.: Designing for diversity: developing complex adaptive tangible products. In: Proceedings of the 1st International Conference on Tangible and embedded interaction, pp. 155–158 (2007)
17. Magerkurth, C., Engelke, T., Memisoglu, M.: Augmenting the virtual domain with physical and social elements. In: Proceedings of the International Conference on Advancements in Computer Entertainment Technology, pp. 163–172 (2004)
18. Montemayor, J., Druin, A., Chipman, L., Farber, A., Guha, M.: Sensing, storytelling, and children: Putting users in control. Technical report, University of Maryland, Computer Science Department (2004)
19. Konkel, M., Leong, V., Ullmer, B., Hu, C.: Tagaboo: A collaborative children's game based on wearable RFID technology. ACM Journal of Ubiquitous Computing 8(5), 382–384 (2004)
20. Itoh, Y., Akinobu, S., Ichida, H., Watanabe, R., Kitamura, Y., Kishino, F.: TSU.MI.KI: stimulating children's creativity and imagination with interactive blocks. In: Proceedings of the 2nd International Conference on Creating, Connecting and Collaborating through Computing 2004, pp. 62–70 (2004)
21. Nakamura, I., Mori, H.: Play and learning in the digital future. IEEE Micro 19(6), 36–42 (1999)
22. Garvey, C.: Play. Harvard Univ. Press, Cambridge (1990)
23. Kaltenbrunner, M., Jordà, S., Geiger, G., Alonso, M.: The reacTable*: A collaborative musical instrument. In: Proceedings of the Workshop on Tangible Interaction in Collaborative Environments (TICE), at the 15th International IEEE Workshops on Enabling Technologies (WETICE 2006), Manchester (2006)
24. Hornecker, E., Buur, J.: Getting a grip on tangible interaction: A framework on physical space and social interaction. In: Proceedings of the ACM International Conference on Computer-Human Interaction (CHI 2006), pp. 437–446 (2006)
25. Gaver, W., Bowers, J., Boucher, A., Gellerson, H., Pennington, S., Schmidt, A., Steed, A., Villars, N., Walker, B.: The drift table: designing for ludic engagement. In: Proceedings of the ACM International Conference on Computer-Human Interaction (CHI 2004), Extended abstracts on human factors in computing systems, pp. 885–900 (2004)