Sound and Interaction for K-12 Mediated Education

David Birchfield*, Thomas Ciufo*, Harvey Thornburg†, Wilhelmina Savenye††

*Arts, Media and Engineering, Arizona State University
dbirchfield@asu.edu, thomas.ciufo@asu.edu
†Arts, Media and Engineering/Fulton School of Engineering, Arizona State University
harvey.thornburg@asu.edu
††College of Education, Arizona State University
willi.savenye@asu.edu

Abstract
Through experiences with developing technologies available on the Internet, in video games, and in new media entertainment, many of our students’ social and intellectual experiences utilize digital media in some fashion. The ever increasing accessibility of sophisticated media tools suggests that such enabling technologies could be ubiquitous in our classrooms and might soon serve as a foundation for learning. However, meaningful integration of interactive digital media is limited by a number of challenges relating to tools, curricula, and evaluation. Our work in K-12 Mediated Education takes an innovative approach to these challenges and utilizes methodologies drawn from interactive computer music for the realization of constructive learning environments that enrich students’ understanding of sound and movement.

1 Introduction

Over the past century a number of educational philosophers have lead a reform movement to reshape the face of learning. Constructivism is a philosophy or approach to learning that posits that learners construct much of what they learn (Schunk 1996). Its proponents emphasize the necessity of play and exploration in self-guided learning, particularly in mediated environments (Jonassen 1999). Instructional teaching can play a complementary role in new learning environments where advanced students and teachers providing feedback create a “zone of proximal development” that speeds learning (Vygotsky 1978). Jonassen and Land hold that education is experiencing a “renaissance in the evolution of learning theory” that is afforded by new technologies (Jonassen and Land 2000). The Theory of Multiple Intelligences (Gardner 1993) posits that there are multiple forms of knowledge, including musical and kinesthetic, that support and reinforce one another. This research corroborates evidence that students have diverse learning styles that are influenced by their life experiences, cultural background, and genetic predisposition. Active learners and visual learners are examples of students who thrive in contexts that extend beyond conventional book and lecture teaching methods. Student Centered Learning Environments (SCLE) encompass a great variety of such learning systems, including open-ended learning environments, microworlds, goal-based scenarios, cognitive apprenticeships, and constructivist learning environments (Land and Hannafin 2000). Such environments have been extensively used for math and science education but rarely for arts and media oriented education (Bransford, Brown et al. 2000).

Interactive digital media affords compelling new opportunities to enrich and extend these proven learning approaches, but it has been underutilized in classrooms to date. In most pre-service teacher training, the scope of media is often restricted to the presentation of audio/visual materials that students passively watch in the classroom. Although these static media resources can serve to vary instructional methods, they do not engage active learners. Educational video games have been introduced in a number of classrooms as a complement to conventional teaching methods (Kafai 1995). Games can be a powerful motivational force in the classroom. However, most games offer only limited modes of engagement and rely on conventional mouse/keyboard/controller interfaces. These interfaces limit the naturally expressive capabilities of students.
In our own related work we have investigated the use of new interfaces to sound (Birchfield 2004; Birchfield, Mattar et al. 2005; Birchfield, Mattar et al. 2005; Ciuflo 2005) that is informed by extensive prior work in this domain. We have looked to additional related work in the development of perceptual and intelligent spaces for performance and interactive installations (Paine 1998; Wren, Basu et al. 1999). This related work suggests that dynamic, interactive sound can engage audiences in truly participatory modes of exploration, but much work is needed to adapt these paradigms to suit the needs of structured learning environments.

In this article, we present our work in K-12 mediated education where we are developing interactive frameworks to engage students’ naturally expressive capabilities and support broad learning objectives. We describe our approaches to curriculum design, interactive media tools, and evaluation. We conclude with a summary of our preliminary findings and an outline of future directions.

2 Learning Objectives

Two overarching learning objectives guide our work: students should develop a (1) deeper understanding of movement and a (2) deeper understanding of sound. These broad objectives are subdivided into demonstrable skills that are taught and evaluated through specialized curricula. We are using a modular approach to curriculum design wherein sets of specialized learning exercises can be grouped to form complete sessions. Multiple sessions can be strung together to form learning paths spanning across weeks or months in a classroom. Our teaching embraces constructivism and experiential learning philosophies and utilizes a number of instructional methods including guided discussion, collaborative exercises, and discovery learning with interactive media.

In this article we focus on our work in developing frameworks for learning through sound and interaction. In this domain we have drawn from Arts Standards at both the regional and national levels (ArtsEdge 2006; Education 2006). The use of arts standards has been a controversial subject in the United States, and indeed a detailed reading would provide an incomplete approach to our students’ learning. However, these standards do provide a broad schematic and well researched framework for the development of skills and knowledge that pertain to a deeper understanding of sound. Furthermore, these arts-based approaches encourage participation, collaboration and active engagement.

We have subdivided our overarching sound learning objective into a set of skills that we are currently building into our curriculum. Specifically, we have designed learning modules and assessment strategies that guide students to:

- distinguish between properties of sound
- reproduce properties of sound using vocalizations
- articulate what they hear using increasingly sophisticated language descriptors
- recognize abstracted sound features from sound environments
- navigate complex sonic environments
- use interactive sound to communicate an idea

3 SMALLab

Central to our work is the development of the Situated Multimedia Arts Learning Lab [SMALLab]. SMALLab is an interactive media environment developed by a team of artists, educators, engineers, media designers and psychologists at the Arts, Media and Engineering program at Arizona State University.

3D Vision based movement tracking, an audio microphone array, a top-mounted video projector and a multi-channel surround audio system comprise an open, yet immersive media environment that is dynamic and interactive. A computing cluster drives the interactive system with custom software for fused multimodal sensing, context modeling, and dynamic visual and sonic feedback. The lab allows teachers and groups of students to interact with one another and their composed media worlds through free play, structured movement, and vocalization. The use of a multimodal sensing apparatus engages the naturally expressive capabilities of K-12 students. Multimodal real time feedback provides a platform for curriculum design that addresses the needs of students with diverse learning styles. SMALLab’s open architecture facilitates social interaction and collaborative knowledge acquisition, emphasizing constructive as opposed to instructive modes of learning. The system is low-cost, re-configurable, and can be transported for easy installation in a classroom or community center.

4 Education Activities

The SMALLab curriculum features a modular design approach, in which teachers may utilize specific combinations of modules which most clearly support their
learning objectives. These learning modules combine varied methods including guided discussion, structured exercises, and open exploration. A variety of educational modules have been implemented and tested with students between eight and twelve years of age. In this section, we describe select modules that are specifically adapted for sound, listening, movement, and real time sonic interaction. We highlight both educational and technical considerations.

**Mediated Listening** relies on guided discussion and is designed to help learners distinguish between properties of sound, more accurately describe features of sound, and abstract these features for later use in more complex sonic environments. We have adapted this exercise from Pauline Oliveros’ deep listening practice (Oliveros 2005) and concepts from the acoustic ecology community (Truax 2001). Students are asked to sit on the floor, and the facilitator leads a discussion about sound with a particular emphasis on distinctions between hearing and listening. The students are then invited to close their eyes, and listen carefully to all the sounds inside and outside of the space. Sound localization is then introduced, and the facilitator rings a small bell while asking the students to point to where the sound is located. After a short period, the sound of the physical bell is replaced by the playback of a recorded bell projected from the multi-channel surround audio system in the SMALLab. The students continue to point to the sound location as the bell sound is moved throughout the space. This focused listening is followed by a discussion of what the students heard, and how their perceptions changed from hearing an acoustic sound, to hearing a recorded sound projected into the space. This exercise can be expanded by listening to, and discussing a variety of sounds, including sounds recorded by the students outside the classroom.

**The Rainstorm** includes collaboration, interaction, listening, and creation. It is a structured exercise that allows learners to mimic sounds through vocalization, collaborate in using sound to communicate an idea, and recognize sounds. In this improvisational exercise, groups of students use their bodies and voices to create the sound of a rainstorm. First the storm is coming in the distance, then it gets closer and becomes more intense until it is right overhead, and then passes into the distance. After listening and creating with only their own sounds, a prerecorded rainstorm is played into the space, and follows a spatial trajectory that the students are trying to create. We have found that the addition of the recorded storm helps the students to be more expressive in their sound making, and helps them more fully engage the enveloping sound of the exercise. This module reinforces the relationship between spatial movement and sound in our understanding of the world.

**The Sound Poem** is the most sophisticated module in our curriculum. It supports free exploration and creation in an interactive sonic environment. It is designed to improve students’ ability to navigate complex sonic environments and use interactive sound to communicate an idea. For this exercise, we have created an interactive environment that fully utilizes multimodal sensing, and real time sonic feedback. The primary goal is to provide students with an opportunity to express themselves through movement and sound, learning the fundamental concepts of sound, gesture and interaction in the process. While most traditional learning environments are physically passive, and don’t allow the student to use their bodies, this exercise encourages the students to construct and shape their sonic environment using full body movement and gesture.

![Figure 3. A student creates a Sound Poem in SMALLab](image-url)

A motion tracking system monitors the position and velocity of a colored ball that the student uses as the primary interface to the system. The tracking system provides three-dimensional position of the object at a rate of 33 fps along with its velocity and acceleration. This data is broadcast to the sound feedback computer, which uses the movement information to control an interactive sound playback environment. The sound engine, built in Max/MSP, contains five sound file players that can be spatialized into two-dimensional locations in the space. The volume of the five players is controlled by the x/y position of the ball in the space, so as the ball gets closer to a given location, the sound assigned to that location also gets louder. These locations are reconfigurable, and the proximity to gain control curves for each sound location can be changed according to a student’s request or curriculum needs. We have developed an extensible database of sound files for this playback system, consisting of both synthetic and recorded sounds. Through interaction with the sounds, students can build individual sets that express a particular idea as they move in the space. One example is the use of environmental sounds related to either specific locations or weather conditions. For instance, we can create a hybrid space combining various wind, rain, or thunder sounds. In addition to the location-based mixing, the velocity of the ball modulates the playback speed and pitch of a given sound. This allows students to drastically manipulate the quality of a given sound by staying in the same general location, but quickly moving the ball. The z-plane location of the ball controls the influence of a band pass filter on
each sound. This allows students to sweep the filter range from low to high as they move the ball from the floor to up over their head. Filter resonance varies directly with the accumulated ball velocity. We have sought to develop sound interaction frameworks that provide immediate, intelligible responses to young student’s behaviors. We have been pleased that the current combination of rich sound sources and multiple degrees of freedom in the students’ movements yields an expressive and engaging learning environment. As students progress through the curriculum, the available resources in both sound and interactive mappings will grow proportionally more sophisticated and malleable. In this regard, educators can gear the complexity of the experience to suit individual students’ progress.

![Figure 4. 3-d trace of a student’s movements](image_url)

The Sound Poem Reflection exercise combines the exploratory model of the Sound Poem with guided discussions from the Mediated Listening exercise. This learning module allows students to develop a deeper understanding of the relationships between sound and movement embedded in their Sound Poems. As students create their Sound Poems, all movement data is archived. After each poem, the facilitator can replay the archived data, and a virtual ball is projected onto the floor of the space. The projected ball reenacts the original movement and initiates the same sonic interactions as in the original. Students reflect on the movements and sounds, now as observers, and discuss the patterns that emerge. To aid in understanding of a longer history of their movements, an accumulated composite trace of their three-dimensional movements (e.g. Figure 4) is also projected on the floor and rotated for multiple viewing angles. The facilitator can lead the students in discussions of their movement and sound interactions, reflecting on similarities and differences across the poems, and articulating various salient features. The orientation of such discussions around the students’ actual sound interactions motivates their learning and facilitates richer understanding.

5 Evaluation

The development of new approaches to learning demands careful attention to appropriate methods of evaluation. Constructive and experiential learning environments are difficult to evaluate given the open ended nature of their methodology. Similarly, conventional arts education frameworks resist quantitative measures. To evaluate our work in mediated education we are developing a multifaceted approach that draws from education, psychology, and computation. We are designing assessment frameworks that provide a comprehensive view of students’ progress, evaluate the success of these new education tools, and examine teacher performance.

Empirical observation plays an important role in assessment. In-person observation occurs first during learning sessions with students. Furthermore, we have developed the Edulink website (Birchfield, Savenye et al. 2006), an online portal for observation and annotation. This site serves as a managed repository for video documentation of each learning session. Here, our network of researchers, educators, media designers, and teachers can view documentation and submit annotations of their observations. Over the course of our pilot studies we have collected a diverse set of empirical observations that assess the effectiveness of all aspects of our work.

We also employ psychological methodologies to assess learner progress. We have designed perception/action studies that evaluate students’ ability to navigate multimodal environments, and assess the impact of interactive sound in visual and movement perception tasks. We have also designed written questionnaires that are administered immediately after a learning session. These questionnaires directly probe students’ understanding of topics including naïve physics, movement/sound relationships, and instructional methods.

As a complement to these known methods, we are developing innovative paradigms for embedded evaluation that utilize computational models of context and knowledge representation. We are developing models to evaluate aspects of students’ sound and movement interactions: first, to extract the context of interaction, including elements and principles underlying the artifacts of each student’s interaction; second, to track how this context develops over time and is individuated amongst different students. The context consists of lexica representing the students’ modes of interaction within the space, plus a syntax describing how lexica relate across different times or modalities. For example, what are the patterns in sound that emerge from the relationship of 3D position and the dynamics of movement during various learning modules? Over time, we expect that aspects of the interaction syntax will become highly consistent within a given exercise, but should include distinct elements across different exercises and across diverse student groups. For example, over time we expect that a given student’s Sound Poem will contain some
distinct, yet consistent movements and sounds. As a student’s learning increases, their poems should diverge from their peers’ in quantifiable ways.

6 Preliminary Results

As of this writing we are undertaking pilot studies with middle school students who participate in after school programs on the campus of Arizona State University under the auspices of the Herberger College for Kids. This initial study has focused on a small set of learners, but has yielded encouraging results that have already lead to revision and improvements in our methods.

Firstly, we observe that students are highly motivated to participate in these learning exercises. Often, they are initially shy and reserved in their interactions, but through the introduction of increasingly active movement and sound exercises, students lose their inhibitions and grow noticeably more engaged and expressive in their use of dynamic sounds and movements.

Secondly, students report that they enjoy the social and collaborative aspects of SMALLab. An important design goal has been to develop a situated learning environment that supports human-to-human interaction and promotes collaborative learning exercises. While many screen/mouse/keyboard interfaces can isolate students and limit their creative capabilities, we are encouraged that here, students are able to engage with their peers in a meaningful way.

Thirdly, we can see that our use of interactive sound is having an impact on students. For example, we expected that the interactive sound design utilized in the Sound Poem engine should bias students to move in particular ways. Specifically, the topology of virtual sound locations should reinforce 2D spatial movements toward the four corners of the space. The interactive relationships between z-plane movements and object velocity should encourage shaking and swooping physical gestures. Indeed, analysis of archived movement data reveals that these similarities emerge for individuals and across groups of students.

7 Further Directions

We have thus far described our research in sound and interaction for a student centered learning environment, SMALLab. Our empirical study and preliminary evaluation have highlighted two critical areas where we are focusing recent efforts: (1) creating an interface for student design of interactivity and (2) expanding our multimodal sensing framework to recognize sonic gestures in students’ collected media and vocalizations.

Section 4 describes a sequence of learning exercises that culminate in the creation of Sound Poems. To date, students have participated in the construction of individual sound environments through two means. First, they can record and select the soundfiles that are used in their poems.

Second, they can specify the spatial location of these sounds in the interactive field. Our work with young students has shown that this limited scope is suitable for novices. However, as students grow in their sophistication with sound and interaction, we must provide increasingly malleable tools for them to create idiosyncratic and expressive interactive environments. To this end, we are currently developing a high-level, Java-based interaction design environment that sits atop a library of custom sonic and visual render engines written in Max/MSP/Jitter. Within this environment, users can design and deploy interactive media environments that are appropriate for particular learning exercises and express their individual understanding of sound and gesture. Our future work will extend this approach to include semi-automated approaches to defining interactive mappings that are both context sensitive and user adaptive.

In related work, we are developing additional modules that promote students’ awareness of their sonic environment and their ability to recognize abstracted sound features therein. We have designed a Sound Harvesting learning module where students explore sound environments outside the classroom with portable audio recording devices. Students are guided to identify and capture sounds that exhibit distinct contours across elementary sound attributes such as amplitude, pitch, and timbre. Recorded sounds are manually segmented, stored in a database, and twice annotated: once by a developing automatic feature-based analysis and sonic gesture recognition engine (which will be discussed shortly), and once by students’ subjective classification. We are designing follow-up exercises lead by teachers in SMALLab where students can retrieve individual sounds or classes of sounds by generating related gestures with their voices. This active exploration and experimentation will lead to informed discussions of the elementary properties of sound as students will better understand the quantitative presence and qualitative meaning of these attributes.

Our current sonic gesture recognition engine is focused on the identification of up-down pitch gestures, but these feature extraction and gesture recognition modules are easily generalized. Pitch feature extraction uses the maximum-likelihood method of (Thornburg and Leistikow 2005). For each STFT frame, the maximized likelihood is thresholded to obtain a binary confidence (pitch/non-pitch indicator) for that frame. These confidence and pitch estimates form the observation layer of a dynamic Bayesian network (DBN). The DBN encodes a state that contains the inherent pitch value for the gesture. The state’s evolution is controlled by a mode that is a discrete variable describing the current activity for the gesture and containing occlusion hypotheses that use confidence information from the Thornburg-Leistikow pitch estimate. Building up from the pitch detection algorithm in this manner, we can construct a mode-transition distribution, or stochastic grammar, that describes the contours of a given gesture. For example, a
given sound gesture can be represented by a sequence of mode values that represent up; down; up/pitch occluded; down/pitch occluded; and null trajectories. Similar grammars will be developed for different gestures and audio features as this work continues. Currently, the gesture segmentation and inherent-pitch trajectories are computed in real time at a rate of 20 frames per second using a standard Gibbs particle filter (Carter and Kohn 1994). Preliminary results show considerable robustness to interference-based occlusions resulting from high noise levels during typical learning activities in SMALLab.

8 Conclusions

We have described our work with interactive movement and sound for experiential learning that includes the realization of SMALLab, a specialized modular curriculum, and a multifaceted approach to evaluation. We have reported on initial findings and we are encouraged by feedback we have received from students, teachers, and parents regarding their learning experiences.

Our future work is directed toward ambitious expansion of each aspect of the project. We will broaden the scope and reach of the learning modules and develop necessary SMALLab software infrastructure to support its implementation. We will work to increase the depth of our computational models of evaluation. Finally, we will expand our outreach activities to deliver this work to a large population of students by embedding SMALLab in several classrooms in our region and nationally.

9 Acknowledgments

We gratefully acknowledge that this work is supported by the National Science Foundation CISE Infrastructure grant under Grant No. 0403428 and IGERT Grant No. 0504647.

10 Further Documentation

Extensive media documentation of this work can be found at: http://ame2.asu.edu/projects/ameed.

References


