Gestural Musical Affordances

Atau Tanaka
Goldsmiths Digital Studios
Goldsmiths, University of London
London SE14 6NW, UK
a.tanaka@gold.ac.uk

Alessandro Altavilla
Culture Lab
Newcastle University
Newcastle upon Tyne NE1 7RU, UK
a.altavilla@gmail.com

Neal Spowage
MTI Research Centre
De Montfort University
Leicester LE1 9BH, UK
nealunreal@gmail.com

ABSTRACT

This paper is a comparative study of gestural interaction with musical sound, designed to gain insight into the notion of musical affordance on interactive music systems. We conducted an interview base user study trialing three accelerometer based devices, an iPhone, a Wii-mote, and an Axivity Wax prototype, with four kinds of musical sound, including percussion, stringed instruments, and voice recordings. The accelerometers from the devices were mapped to computer based sound synthesis parameters. By using consistent mappings across different source sounds, and performing them from the three different devices, users experienced forms of physical, sonic, and cultural affordance, that combine to form what we term musical affordance.

1. INTRODUCTION

1.1 Gestural control of sound

The live performance of computer based music has become completely democratized. Research in this area capturing performer gesture through sensors, and mapping them to articulate computer synthesized sound is represented by the New Interfaces for Musical Expression (NIME) field. Meanwhile consumer products are widely available, with mobile phone apps simulating traditional instruments like guitar and woodwinds, played on the touchscreen and making use of the tilt sensors of the mobile phone. Nintendo's Wii Music transforms their Wii-mote accelerometer-based game controller to control on-screen virtual musical instruments. Can these devices be considered musical instruments? Is there a new type of musical instrument, a NIME instrument, that is an expressive gesture sensing hardware/software system? In order to answer this question from an interaction design perspective, we apply established notions of affordance to simple gestural interactive music systems in a laboratory context. This paper presents a comparative study of three gestural controllers (an iPhone, Wii-mote, and Axivity Wax prototype) controlling a bank of different computer based sounds. Through a series of interviews, we gain insight into users trying these systems, and map out a complex relationship of physical affordance, sonic affordance, and cultural association that combine to make the total musical affordance of an interactive music instrument.

1.2 Affordance

1.2.1 Affordance definition

Affordance is a well established concept in interaction design that describes what in the physical attributes of an object can be compatible with those of an actor and invite or make possible an action on that object. Affordances are a configuration of properties that provide a direct link between perception and action.

1.2.2 Affordances in music interfaces

Sound is one of the physical properties that can contribute to the affordance provided by an object. This has been studied in a range of related disciplines from auditory perception to music software design. Gaver's seminal work in this field focuses on non-musical, or everyday sounds [13]. With this the focus is on the use of sonic events in computer interfaces.

> *In distinguishing everyday and musical sounds, the emphasis should be on the kind of experience a given sound affords, whether it is one of the sound itself or of its source. In terms of the interface, the distinction is one of the dimensions of sound that are used to represent data.*

Brazil and Fernstrom extend this work on auditory icons in a volume dedicated to data sonification [4].

In music, affordances have been used by Magnusson [18] to inform the design of "screen-based interfaces as digital musical instruments." Network music collaboration has been described in terms of affordance by Gurevich [16], Brassch [2], and Dillon and Brown [10]. The musical affordance of synthesis parameter mappings on mobile devices has begun to be described by Tanaka [22]. The musical affordance of everyday objects has been explored by Cook and Pullin [8]. The concept of affordance has also been used in the sociology of music by DeNora [9] to describe the sociality of certain pieces of music.

The integration of sound as an affordance-giving property is consistent with the ecological origins of the concept. Audio becomes an additional modality through which complex affordances – sequential, nested, and multimodal – can be conveyed.

> *Sound conveys information for an affordance which can not be seen. (Gaver)*

The use of affordance to describe interactions with musical processes and with works of music is more problematic as it necessarily demands engagement with cognitive and cultural processes.

1.2.3 Culture in Affordance

The role of culture in affordance is contested. According to Gibson's original ecological conception of the term [14], the environment and its properties offer enough information, that when perceived by an actor, provide indications for action without intermediate stages of memory or inference. Cognitive approaches emphasize that people have access to sensations, memories, and build up symbolic representations which guide their actions in the environment around them. Norman's application of affordances in design practice includes cultural and social constraints - affordance is shaped by person's previous knowledge and experience [19].

The context of use, whether it is cultural or not, can have an effect on how the functional affordance of physical properties of an object are perceived. Oshlyansky in an almost humorous experiment demonstrates that cross-cultural differences can invert the perceived affordance of a simple light switch [20]. Rizzo coins the term *intentional affordances* to describe the production of affordance in action and evoked response over more than a lifetime, pointing to a neurological basis for cultural learning [21]. He cites cognitive scientist Tomasello in looking artifacts, or cultural objects, and intentional relations between actors,

> *Such affordances rest upon the understanding of the intentional relations that other persons have with that object or artefact—-that is, the intentional relations that other person have to the world through the artefact. [23]*

This points to the notion of *tailoring culture* that has been used by Carter and Henderson [6], Maclean, Carter, Lovstrand and Moran [17], and Won, Steimerling and Wulf [24] in software design.

1.2.4 Learning and memory in auditory affordance

In the auditory affordances of sonic icons, warnings, and alerts, Brazil and Fernström [4] note that these are things that do not come naturally, but are "something that is learned in a social and cultural context"

In this paper we recognize the complexity of cultural contexts in pinning down affordances in interactive music systems. We present a comparative study of gestural interaction with sound on three similar devices - two commonly available consumer products, and one research prototype. With the experiment, we seek to identify physical affordances of the object, the affordance of sound articulated by user interaction with these gestures, and through a combination of physical and sonic affordance, arrive at a proposed notion of what constitutes musical affordance.

2. METHOD

2.1 User study

We conducted a comparative user study of three small devices that enabled gestural control of computer based sound production. These devices were: a Nintendo Wii-mote, an Apple iPhone, and an Axivity Wax prototype [1]. These objects are similar in that they use accelerometer sensors to detect tilt, rotation, and movement of the device. While they are all small, wireless devices, they have very different form factors. They also have different default contexts that are associated with their customary use. In these studies, these three devices transmitted their accelerometer data wirelessly to a host computer running a series of sound synthesis patches. The gesture of the user on each device was picked up by the accelerometer and controlled the triggering, and time/frequency modulation of a range of individual digital sound samples. We used four sounds: a triggered snare drum, a single violin note, a continuous violin tone, and a looped spoken vocal phrase. By comparing how users interacted with the three objects in articulating these sound through arm gesture, we were interested to look at the physical affordances of these objects. By programming different modes of sound interaction, we were interested to see if the sounds themselves offered a form of sonic affordance.

The study took place on two UK universities, one day at each location. The participants were graduate students and university staff as well as members of the general public in the community surrounding the university. The three objects were placed on a table in a neutral room and high quality stereo loudspeakers centered along one wall, giving uniform sound distribution. The study subject stood in front of the table, and an interviewer in the corner by the laptop.

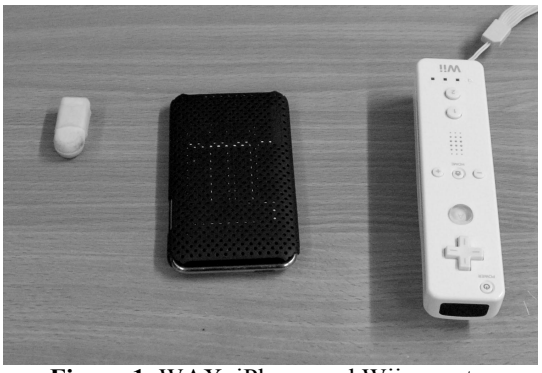

Figure 1. WAX, iPhone and Wii-remote

The subjects were allowed to choose the order in which they selected the devices. They were placed in order left to right in front of them. From their perspective this was Wii-remote, iPhone with covered screen, and WAX (Figure 1). The order of initial choice allowed the users to develop and justify a favourite device.

The study was filmed with two synchronized video camera giving a front view of the participant and a diagonal detailed view of the participant's arm and hands as they manipulated the objects. A microphone recorded sound

from the study and interview onto the audio track of one of the video cameras.
The host computer running the sound synthesis patches connected to the loudspeakers was in a discreet place or invisible to the participant. The laptop was in the corner of the room with its screen turned away from the participant Newcastle and behind a two way mirror in a control room in Leicester.
Each session with a participant lasted between 20-30 minutes. The study was structured into three main sections. 1.) A pre interview to ascertain the musical and technical background of each participant, 2.) Interactions with the controllers and sound, and 3.) An interview.
There were sixteen (16) participants, nine (9) in the Leicester study and seven (7) in the Newcastle study, aged 23-48. There were five female participants and eleven male participants.

2.2 Four Interactions

The user experience was designed to achieve a progressive grade of interaction, based exclusively on the movement of the controllers, between the users and the system, starting from a simple trigger mechanism to the control of a complex voice sound, listed as above:
a) The triggering of a single, non changing, drum snare sample;
b) The triggering of a single, non changing, sound of a bowed violin;
c) The control of a granulator, looping the sample of a violin chord. Modification of pitches and speed were possible due to the movement of the controller on the 2 axis (X, Y);
d) The control of a granulator, looping the sample of a voice sample; modification of pitches and speed were possible due to the movement of the controller on the 2 axis (X, Y);
The drum snare sound was chosen for the cultural association with the act of drumming. On the second and third tests we chose the classical violin sound, an original recording of a bowed A and E string, because the instrument is well known in western culture, and because it is a simple sound with a rich timbre. The fourth test, a voice sample, was a recording of a well known BBC Radio host.
In sounds 1 and 2, sound was triggered a simple threshold trigger on the accelerometer data. We were interested to see if this sound afforded hitting gestures on the part of the user. While the mode of interaction was exactly the same as in the drum we wanted to see if the sound might afford a different gesture, that of mimicking the bowing of a violin. Technically both types of gestures, hitting, or bowing, would have triggered the thresholds in each sound synthesis patch.
Sounds 3 and 4, the continuous violin and voice, were continuously playing, looping sounds that could be sculpted by the two dimensional tilt of the device. To raise the pitch of the sample the user would tip the object on its y axis so that its front end (the point furthest from the user) would rise. This would also allow the user to raise the pitch by holding the object rigid and raising their arm. The sonic effect would be the same. To lower the pitch the user would tip the object so that its front end would drop, or they could hold it rigid and drop their arm. To change the duration of the sample the user would roll the object on the x axis; anti-clockwise to extend duration and clockwise to shorten duration, in relation to its orientation to the user. The operation of the three different hardware devices, in reference to their triggering thresholds, smoothing and pitch/time ranges, was made as consistent as possible. Differences were negligible.

2.3 Interview

There were 19 interview questions. Incoming questions asked about the participant's prior musical experience, whether they had played videogames on a Wii, and whether they used smartphones. Questions about the interactions asked them to describe their favorite device for each sound, and their favorite device/sound combination. Outgoing questions asked about their experience as a whole, whether it was unexpected, and whether it was interesting for them.

3. RESULTS

The interviews were conducted as each participant tried the different device/sound combinations. A series of outgoing questions at the end focused on the whole study and their experience.

3.1 Sound and object preferences

In this process, participants were asked which of the four test sounds (drum, violin note trigger, continuous violin tone, voice loop) was their favorite sound. Figure 2 charts the distribution of favorite sound. Interestingly, the participants tended to prefer the complex sounds that we feared may be more difficult to perform.

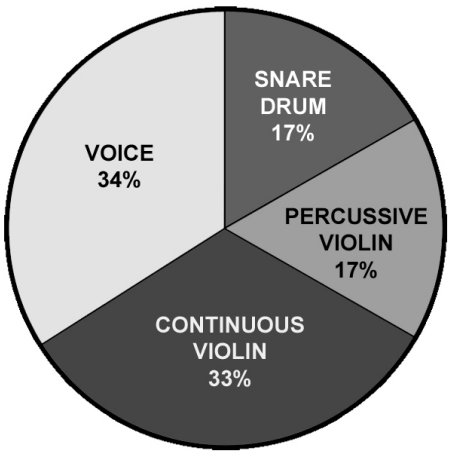

Figure 2. Distribution of participants' favorite sound

As a follow up they were asked for their favorite sound/object combination. While the two complex continuous sounds (violin tone and voice) emerge again, it is curious that a different device is associated with these sounds, despite the fact that the accelerometer parameter mapping was identical. It seemed that some aspect, either of physical affordance of the objects, or the sonic affor-

dance of the sounds were having an effect on the participants' degree of satisfaction.

3.2 Interview results

The interview responses can be divided into several categories:

- Describe an object's physical affordance in relation to gestural sound making
- Compare the physical affordances of the objects
- Describe the sonic affordances of the test sounds
- Compare the sonic affordances of the sounds
- Describe relationship between object and sound
- Describe experience of producing sounds interactively through gesture

3.3 Physical affordances

The quotes are noted with the participant giving the answer (U) and the interview question to which they answered (Q). So "U14Q3" is an excerpt from User 14's answer to Question 3.

3.3.1 Wii-mote

The Nintendo Wii-mote controller was the most clearly described by participants in terms of its form factor and the kinds of gesture it afforded.

This one (Wii-mote) feels slightly more like bulky and restrictive in terms of kind of movement (U9) I still find the Wii-mote quite bulky in terms of gestural motions

It had a clear association as the object as it is known in consumer culture, a video game controller.

I felt awkward holding the Wii, it's a awkward shape, and I think because it's a Wii remote you automatically want to move it the way you would if you'd used a Wii, would be used to using the Wii remote. (U5Q17)

It also evoked other objects, some of them musical, affording certain musical gestures:

The Wii is kind of can remind of a stick and drums, this can have pads, but, well this might, might be better for (laugh), well, it looks more like a stick you can use it like a stick (U14 Q4)

At the cultural extreme, the form factor of the Wii was strong enough to make the user forget its default association as a video game controller to remind them of another everyday media object.

This is one is heavier I think (Wii-mote), and also because it feels like a TV remote…and I don't like TV… (U10Q16)

3.3.2 iPhone

The iPhone's form factor also afforded, according to User 5, a certain ease of use in sculpting the voice sound.

It was easier to play with the different ways that it changed the sound depending on what I was doing. I think as well because it's the shape and size of it, it's a lot easier to move around than either of these (other) two. It's probably the reason I picked it (unintelligible) the other ones as well. It's much more tactile. (U5Q10)

Curiously the iPhone also reminded users of a television remote

I think iPhone wanted me to, like, be, like er, no, like a remote, you know, from television. He wanted be, he wanted too much space on my hand (U14Q9)

We wonder if, once the object is divorced from the activity with which it is usually associated and applied in a new setting of control, that the television remote control becomes the simplest control device from which the user in some way seeks to elicit cultural affordance.

In the last quote, it is interesting to note the use of the word "want". We can interpret this as one way a layperson seeks intuitively to articulate the notion of affordance.

3.3.3 Wax

The Wax was the smallest device with the most nondescript packaging and form – a pill-shaped object about the size of a fingernail. It was also the device unknown to the users, so carried no prior association with it. This allowed the object, physically and culturally, to afford a kind of liberty:

It (the Wax) was smallest and it, it er, gave me more freedom (U14Q7)

It's the simplest one, I find that one the easiest to figure out and the easiest to control. (U4Q9)

When I had this (the Wax) on my hand I didn't have preconceptions about the instrument or the controller or anything linked to this. (U13Q8)

You weren't bound to sort of… being dictated how you might use it. Slightly less than the other two p'raps. (U15Q7)

This freedom translated to responsiveness, clarity and detail, and allowed the participants to focus on the sound

It seems more immediate (U13Q7)

Because it's small I can see in a clear way, mmm, in what direction I am tilting in. (U10)

It feels really nice to handle in. it feels you can get… because it has small surface…you can kind of get more detail with your gestures… (U9)

I didn't really notice that I'm having something in my hand, and it was quite good because I could focus on the pitch (U14Q9)

However questions of scale, one of the key criteria for affordance to exist between object and actor, emerge with the Wax's small form factor.

The wax is interesting because it's so small, but for using only with hands is too small to be comfortable. (U6Q10)

3.3.4 Comparing devices

For User 6, the Wii-mote afforded the gestures necessary to percussive sounds but less so for the violin note.

In terms of shape and holding the Wii is definitely better (for the drum trigger task). (U6Q7)

The Wii-mote is easier to hold but in this particular case (for the violin trigger) it wasn't easy to figure out what it was doing. (U6Q9)

This is very interesting because the triggering interaction was exactly the same for the drum and triggered violin note – there was only a change of source audio sample in the two examples. This points to a difference in sonic affordance that changes the perception of interaction to the user.

Despite identical mappings from x-y tilt to time-frequency modulation of the voice sample, User 4 found one device more "natural" than the other

The Wii-mote it feels unnatural to change the sound with it, on that, that one, but that one felt better (points at iPhone). (U4Q10)

For an explanation we recall User 5's similar answer to the same question:

I think as well because it's the shape and size of it (iPhone), it's a lot easier to move around than either of these (other) two. (U5Q10)

One user compares the cultural association of the iPhone as an iconic telephone compared to the Wax which has no association.

These controllers look, this reminds me (the iPhone) of the mobile communication. But this one (the wax) is not really linked to my mind to anything. (U13Q8)

This is corroborated by User 8's answer to the same question. His word choice recalls the way User 14 intuitively used the word, "want" to describe affordance – User 8 uses the words "force" and "encourage" to describe how these objects afford playing the musical activities in a certain fashion:

I wasn't forced to interact in a specific way (with the Wax), whereas the phone and the Wii-mote encouraged in interacting in a specific way, whereas, this (the Wax) is an object you can play with in which way you want . (U8Q8)

User 8 goes on to develop this line of thought more specifically for 1-D or 2-D interaction:

Because it (the iPhone) was a bit larger than this one (Wax) so it gave me more kind of control over the multiple axis, while the small object (Wax) it's fine to tilt in one direction, to control in one way, but when you want to control two axis is a bit easier with a larger object (like the iPhone). (U8Q9)

Finally this user explains trying to reconcile these qualities, arriving at a comparative description of the three devices that recognizes the common interaction mapping:

Because it (iPhone) gave me a lot of control within the two different axis. I've tried when I was using the Wax to hold on the palm of my hand so I could use my hand as a larger object to rotate, to imitate the iPhone. But I think because the iPhone has a flat surface it allows me to tilt it and move it, to control the axis, while the Wii-mote I guess is more square, a long shape, it emphasizes one axis, the tilting, but not the second (the rolling). (U8Q10)

3.4 Sonic Affordances

One of the hypotheses at the outset of the study was that recognizable musical instrument sounds would afford gestures that imitated actual musical practice on those instruments. This is a phenomenon described by Godøy as "sound tracing" [15]. The drum example did evoke such a response in several participants (cf. video). The violin sound afforded this gesture to only one participant. He describes this sense of affordance developing out of imagining the device attached to the instrument the sound refers to.

When I picked up the iPhone I didn't instantly think maybe I can try it as a violin bow, … I kind found easy to imagine the actual instrument, rather than I was just playing a sample, even if it still didn't really work in that way, but I can imagine it (the iPhone) stuck like on a violin bow maybe. (U13Q7)

Other users admitted that it simply didn't occur to them, or that their lack of knowledge on how to play a violin kept them from trying that gesture.

I recognized it (the violin) but I didn't really think about it (when playing) (U11Q10)

I didn't know enough about the violin to know what you can control effectively. (U7Q16)

3.4.1 The richness of vocal sound

The sound of the voice is not associated with an external physical object; rather it is associated with the human body. Our hypothesis was that this would be an evocative sound that did not have any associations with a physical affordance. Instead, we wondered if people would make gesticulation as if they were talking the phrase.

The participants were drawn to the sound, and describe its timbral and semantic richness, but do not articulate any gesture or movement to associate with that richness.

Errm, it (the voice sound) felt like you could draw more of a variety of textures, p'raps a bit of something, erm, it just, it was a bit, it was a bit more going on it was a bit more interesting to, to use. (U15 Q9)

One user likened the time/frequency manipulation to scratching a turntable.

It was kind of like, scratching I suppose (U1210)

More than one user were interested in the balance between the recognisable power of the voice and its sonic abstraction

I like the kind of fluxuations between being, erm, recognizable as a voice and, recognizable only as sound and, something in between those (U11Q15)

There is a lot of texture variation within the different worlds and I can control them. (U9Q11)

Finally, one user was drawn to the comprehensibility of the vocal phrase

I was trying to understand what was being said. I don't understand the dialect or what language it was in, but I was able to get it in what I thought was in time. But I, rather than speed it up or slow it down. I wanted it to be at the correct speed of what felt was a language being spoken. (U3Q10)

3.4.2 Comparing sounds

The question of abstraction, be it sonic abstraction or a detachment from the source instrument, arises in comparing the different sound examples.

Because it (the voice) is kind of abstract, so it goes the way you are controlling it, while the violin one is a bit strange because you are not really playing a violin. (U6Q11 comparing voice and violin)

3.5 Exploration

As affordance leads to forms of embodied knowledge, exploration is a key part of the interaction between environment and actor. Gaver uses this as a way to frame complex actions, going from passive perception of the environment to its exploration,

"Exploration of afforded actions leads to discovery of the system, rather than knowledge of the system metaphor leading to expectations of its affordances." [13]

For User 4, the sound helped in beginning to understand an unfamiliar system beginning a process of exploration.

I didn't know what it was to begin with and it was quite interesting when it made the noise, and then it was easy to figure out what to do to make the noise. (U4Q7)

The notion of surprise and the unexpected was part of that process for this user.

I like the unexpected sound of a violin, the classical instrument coming out of a piece of modern technology. (U4Q16)

The actual lack of object affordance in the Wax leads to a form of curiosity for another user as an entry point to exploration.

(Points at Wax) Maybe because it's seemingly such an uninspiring little object you can actually, er, create a sound that was um, quite intriguing. (U3Q9)

User 13 answers outgoing questions on the experience as whole and possible future ideas, and evokes the notion of exploring as a way to play a possible dynamic, time varying system.

You could explore the possibility over the time rather than just playing like specific sounds like, I don't know, an instrument you already know, like a guitar or something. (U13Q14)

He goes on to specifically use the word, "exploration" to describe a process of finding constraints and forms of negative reinforcement to define the instrument.

It's like an exploration, like just finding out, sort of testing it. Seeing what the boundaries, or where the limitations were, see if I can get something nice, in terms of movement. To see what didn't work. (U13Q15)

Another user also indicates a desire of an ongoing process to retain interest, not dissimilar to the previous user's time varying idea. This supports Gaver's opposition of knowledge and expectation of affordances.

I like the exploration but also, er, a little bit limiting once you realized what it could do, or possibly not limiting it would be nice to spend a lot longer with, each object cuz it felt like, obviously this is a test but, once I understood how, how it was working, I wanted to continue playing around (U11Q15)

Finally, a user describes the experience of participating in the study as a form of exploration and slowly unfolding understanding.

They (modes of interaction) were invisible at the beginning, but then through experimentation and finding out the limits to each one, and um, it felt more, er, like I had more control towards the end once I'd made the limits of each sensor, the sort of like, visible to me. (U12Q12)

In the videos, we observed other users going through a trajectory of discovery. User 8 was at first frustrated, not being able to trigger the simple drum sound. The two modes of interaction with the violin led to a discovery allowing him to over come his initial frustration, and the voice example was described by him as,

Fun… really intuitive… enabled me to control the sound in a very direct way. (U8Q17)

3.6 In summary

3.6.1 Physical Affordances

The Wii-remote is restrictive. Better suited to a device that restricts user actions. Hitting and pointing. Rolling gesture was poor with all sounds.

The iPhone is an all rounder conversely it still gave affordance information to the user.

Both Wii-remote and iPhone retained their apparent cultural affordance. i.e. TV remote.

The WAX, with minimal affordance, suggested freedom and responsiveness to the user, conversely the lack of affordance information created one or two attempts to "fit" it into something that had affordance.

3.6.2 Sonic Affordances

The only sounds that afforded an expected gesture, that of hitting, was the sounds of the snare drum and percussive violin.

The lack of presence of a physical violin prevented users mimicking the action of a bowed violin with all devices.

The complexity of the vocal phrase drew the users interest and had no consistent gestural action in conjunction with all devices.

3.6.3 Exploration

Most users were unfamiliar with the devices as sound controllers. This invoked exploration and interest for

each device. In general, they wished to compare and find the limits of each device and the sound it was controlling. Apparently, the more complex the sound that was being controlled, such as the vocal phrase and bowed violin, the more exploration was desired from the user.

4. DISCUSSION

Affordance was originally conceived to describe relationships between the environment and an actor. These relationships are based on perception, complementarity between environment and actor, and are subject to relative scale. When applied to design of objects, affordance extended the notion of actions an environment could offer an actor, to the kinds of use that an object invited a subject. When extended to computer interfaces, this ecological and design knowledge is brought to bear in implementing metaphors for task completion by users of a system. In all these cases, the assumption is that we are dealing with unitary entities – single objects, or interface items with a given function. Complex relationships are described in terms of a sequence of a number of unitary affordances, or the hierarchical nesting of affordances within one another.

With a programmable hardware/software system, the complexity of the entity (in this case an experimental digital musical instrument) and the gestural affordance it offers the user becomes much more difficult to identify. The results of the user study point to this complexity. The relationship between "environment" and "actor" go beyond complementarity and scale, and might begin to include historical information about the actor or the instrument such as musical experience and subjective matters such as musical preference and taste. In this, culture enters into the possible affordance provided by the system. The question arises not whether culture should be part of affordance, but how one might study subjective variation across a range of users rigorously, or how one might design to accommodate evolving cultures. Conversely potential problems arise of designing, or over designing prescriptive systems that lead users to "expect" certain affordances instead of being allowed to explore system characteristics.

4.1 Scale

The Wax may be too small a scale to have gestural affordance to humans. One user tried fixing it to his palm to give it a larger affordance providing surface.

The scale of affordance with these instruments lies not just in the relative sizes of object and user, but of the scale of meaningful movement afforded by the form factor and default cultural association. The Wii-mote gestures tend to be bigger than with the other objects, belying its cultural origins as a controller for playing video games that use a sports metaphor.

The iPhone offered different degrees of gestural interaction, from small-scale movements, such as rotation, tilt, towards big complex gestures involving the whole use of the body and greater production of movements of arms and hands. In this sense, it seemed to offer a broader grade of physical affordance. This observation appears to concur the preferences of the majority of users who matched the iPhone with the complicated and dynamic voice sample rather than the simple violin sample. The smaller Wax, often referred to as a pencil eraser, afforded gestures of intricate examination or drawing in the air. Why would these gestures be generated from an object whose affordance was minimal?

4.2 Music apps and systems integration

The instruments we used in our study reflect the kind of sophisticated system that is increasingly available as studio tools in the music technology field, or as consumer oriented music apps on mobile platforms. One of the best-selling iPhone music apps is the Ocarina [25], a virtual wind instrument where the player blows into the phone's microphone and fingers notes on the multi-touch screen to simulate playing a simple bore-hole woodwind instrument. Does such a mobile phone based virtual instrument afford playing music according to the performance practice borrowed from an acoustic instrument? The results of this study suggest that it could be possible, but that the answer to the question depends not just on the affordance proposed by the interface design, but the inherent object affordance of the iPhone and the latent sonic affordance of a computer generated ocarina sound. This puts three forms of affordance in play in a complex interdependent relationship that is neither sequential nor nested. The complex relationship of these multiple affordances is compounded in that their relative "affordance mix" is dependent on characteristics of the actor, its player.

4.3 Artifacts

The building blocks for these high level instruments are each already artifacts embedded in contemporary culture. Whether it is a consumer device like a video game controller, or a form of collective memory such as a historical instrument (people might be familiar with an ocarina sound without ever having seen or played one), these constructs are cultural objects. While the early affordance literature proposed using the notion of affordance to better design new artifacts [13], the current proliferation of highly charged cultural objects in society might indicate the interest of studying the affordances of component artifacts that contribute to the overall make up of a higher level system. Fleming's model for classical artifact study [11], which creates an iterative structure of artifact makeup, identity, comparison, values, and interpretation to situate historical artifacts in evolving cultures may be a useful technique with which to study the complex relationships of multiple affordances that come together in a complex, compound artifact. Chung and Ishii propose the notion of Mega Affordance Objects to extend primitive objects with multiple functional affordances [7].

5. CONCLUSION

The complex mix of affordances takes place because the components that make up an interactive music instrument

are not neutral materials, but are high level, highly evolved elements each of which brings a baseline affordance in its original home context and culture of use. These high level elements – designing with already-designed components, from mobile phone hardware to sophisticated sound synthesis software libraries, to instrumental metaphors drawing on centuries of music history are the elements which will make up the "environment" with properties we try to characterize in use. These complex components are assembled in a form of systems integration to create highly sophisticated interactive media systems leading to a form of high-level construction that requires an equally high level conception of affordance, not in the unitary singular, but as constructs of affordances in the complex plural.

This is seen in the results of the study where participants' interactions were influenced consistently but not systematically by the physical affordance offered by the form of the device and the cultural affordance projected by the image of instruments like the violin, varying with device and sound despite constant underlying gestural mapping. The gestural affordance of the instrument, then, depends not just on trigger threshold mapping or x-y time/frequency mapping, but will vary by device and source sound, physical affordances and cultural associations.

6. ACKNLOWLEDGEMENTS

We would like to thank the participants of our study.

7. REFERENCES

[1] Axivity http://axivity.com

[2] Braasch, J. The Telematic Music System: Affordances for a New Instrument to Shape the Music of Tomorrow. Contemporary Music Review 28, 4/5 (2009), 421 – 432.

[3] Braun, V., Clarke, V.Using thematic analysis in psychology. Qualitative Research in Psychology, 3:2 (2006), 77-101.

[4] Brazil, E. and Fernström, M. Auditory icons. In Hermann, T., Hunt, A., Neuhoff, J. G., editors. The Sonification Handbook. Logos Publishing House, Berlin, Germany, 2011, 325–338.

[5] Carter, K., Henderson, A.. Tailoring culture. In: Hellman, R., Ruohonen, M., Sorgard, P. (Eds.). In Proceedings of the 13th IRIS, Reports on Computer Science and Mathematics, 107 (1990), 103-116.

[6] Carter, K., Henderson, A., 1990. Tailoring culture. In: Hellman, R., Ruohonen, M., Sorgard, P. (Eds.), Proceedings of the 13th IRIS, Reports on Computer Science and Mathematics, No. 107, Abo Akademi University, pp. 103-116.

[7] Chung, K., Ishii, H. Fusing computation into mega-affordance objects. CHI 2009 Workshop on Transitive Materials. Tangible Media Group, MIT Media Lab (2009).

[8] Cook, A., Pullin, G. Tactophonics: your favourite thing wants to sing. In Proc. NIME07, ACM Press (2007), 285 - 288.

[9] DeNora, T. After Adorno: Rethinking Music Sociology. Cambridge University Press, Cambridge, United Kingdom, 2003.

[10] Dillon, S. C. and Brown, A. R. The educational affordances of generative media in arts education. In Proc. INTED2010. Valencia, Spain (2010).

[11] Fleming, E. McC., Artifact Study: A Proposed Model. *Winterthur Portfolio*, Vol. 9. (1974), pp. 153-173.

[12] Gaver, W. W. The Sonic Finder: An Interface That Uses Auditory Icons. Human-Computer Interaction 4, 1 (1989), 67-94.

[13] Gaver, W. "Technology Affordances" In *Proceedings of the CHI'91 Conference*. 79-84, 1991.

[14] Gibson J. J. The Ecological Approach to Visual Perception. Houghton Mifflin, Boston, USA, 1979.

[15] Godøy, R. "Geometry and Effort in Gestural Renderings of Musical Sound." In Dias, M., Gibet, S., Wanderley, M., Bastos R. (Eds.) *Gesture-Based Human-Computer Interaction and Simulation LNAI 5085*. Springer. 2009.

[16] Gurevich, M. JamSpace: a networked real-time collaborative music environment. Ext. Abstracts CHI 2006, ACM Press (2006), 821-826.

[17] MacLean, A., Carter, K, Lovstrand, L., Moran, T. P. User-tailorable systems: Pressing the issue with buttons. In Proc. CHI 1990, ACM Press (1990), 175 - 182.

[18] Magnusson, T. Affordances and constraints in screen-based musical instruments. In Proc. NordiCHI 2006, ACM Press (2006), 441-444.

[19] Norman, D. A. The Psychology of Everyday Things. Basic Books, New York, 1988.

[20] Oshlyansky, L., Thimbleby, H., Cairns, P. Breaking Affordance: Culture as Context. In Proc. NordiCHI 2004, ACM Press (2004), 81-84.

[21] Rizzo, A. The Origin and Design of Intentional Affordances. In Proc DIS 2006.

[22] Tanaka, A. "Mapping Out Instruments, Affordances, and Mobiles." In *Proc. New Interfaces for Musical Expression (NIME)*. pp. 88-93. 2010.

[23] Tomasello, M..The cultural origins of human cognition. Harvard University Press, Cambridge, MA, USA, 1999.

[24] Won, M., Stiemerling, O., Wulf, V. Component-Based Approaches to Tailorable Systems. In Lieberman, H., and Paterno F., Wulf, V. (Eds.): End User Development. Kluwer Academic Publishers, Netherlands, 2006.

[25] G. Wang, Designing Smule's iPhone Ocarina, in Proc. Intl. Conf. New Interfaces for Musical Expression, Pittsburgh, PA, USA, 2009.