Haptic Wave: A Cross-Modal Interface for Visually Impaired Audio Producers

Atau Tanaka
Goldsmiths, University of London
London, UK
a.tanaka@gold.ac.uk

Adam Parkinson
Goldsmiths, University of London
London, UK
a.parkinson@gold.ac.uk

ABSTRACT
We present the Haptic Wave, a device that allows cross-modal mapping of digital audio to the haptic domain, intended for use by audio producers/engineers with visual impairments. We describe a series of participatory design activities adapted to non-sighted users where the act of prototyping facilitates dialog. A series of workshops scoping user needs, and testing a technology mock up and lo-fidelity prototype fed into the design of a final high-spec prototype. The Haptic Wave was tested in the laboratory, then deployed in real world settings in recording studios and audio production facilities. The cross-modal mapping is kinesthetic and allows the direct manipulation of sound without the translation of an existing visual interface. The research gleans insight into working with users with visual impairments, and transforms perspective to think of them as experts in non-visual interfaces for all users.

Author Keywords
User Centered Design; Multimodal Interaction; Cross modal mapping; Haptic interfaces; Digital audio production

ACM Classification Keywords
H.5.2 User Interfaces: Haptic I/O

INTRODUCTION
Editing recorded audio is a complex task in music production, radio and other broadcast media, and speech transcription. Basic time domain editing is a task where sound bites may be trimmed of extraneous content before and after the moment of interest. In more advanced cases, individual words in a text or notes in a melody may be replaced by "splicing" out erroneous content and inserting a better "take." More difficult tasks that require skill and training involve the identification of silent segments, detection of sound onset, and the precise definition of loops in rhythmic musical time.

The digitization of audio facilitated audio editing by providing random access to content, obviating the need to wait for a tape to rewind or fast forward. Computer-based digital audio editing allowed a basic form of multi-modal interaction by providing a visual representation of the audio being edited in the form of a graphic waveform display. This allowed the user to see, as well as hear, the silences, transients, and individual audio events. The visual waveform has become a ubiquitous standard in professional audio and media production, and has been popularized in the graphic user interface of music sharing websites such as SoundCloud[1].

In the analog era, audio editing was performed on open reel tape recorders. The producer would play the tape to approximately the point of the desired edit, stop the motor and manually turn the reels back and forth, "scrubbing" the area of interest across the playback heads, allowing the user to hear in slow motion the edit point. After marking the beginning and end points of the edit with a marker, they would cut the tape with a razor blade. They would then splice together the tape (or insert a new section from another take) with a piece of adhesive. The editing process depended entirely on one sensory modality, that of hearing, and rendered the process tactile and visceral. One can say that analog audio editing took place in a paradigm of direct manipulation of content. Interestingly, the words "scrubbing", "cutting" and "splicing" from the analog era remained as the metaphors as these tasks were revolutionized in the digital era.

Figure 1. A typical visual representation of an audio waveform.

The visual waveform display allows the user to "see the sound" (Fig. 1). However, these advantages of graphical audio editing are not accessible to users with visual impairments. If cross-modal mapping allows us to substitute one sensory modality for another, could we map the visual aspects of digital audio editing to another sensory modality? Or,

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

CHI'16, May 07–12, 2016, San Jose, CA, USA

ACM 978-1-4503-3362-7/16/05...$15.00.
DOI: http://dx.doi.org/10.1145/2858036.2858304

given that the graphical waveform display is already a cross-modal representation of sound, can we find another cross-modal mapping to a non-visual sensory modality that would facilitate blind and partially sighted users in carrying out digital audio editing? If visual waveform displays allow sighted users to "see the sound", could we build an alternative interface for visually impaired users to "feel the sound"? This paper reports on the design of an interface that allows digital audio editing in the haptic domain.

RELATED WORK

Cross modal interaction

Intermodal perception (where information is available on more than one sensory system), and cross-modal integration (the transfer of cognition from one sensory modality to another) are abilities that are both innate and learned, and are observed early in infant development [34]. Multimodal interaction in HCI takes advantage of intermodal perception to enhance the use of interactive systems by exploiting multiple modalities on the same task [26, 21]. In cross modal integration, sensory substitution takes place where one perceptual modality stimulates another. This can be exploited in the design of interactive systems to aid users of varying abilities [4]. Auditory perception has been used to allow engineers who are blind and partially sighted to navigate data structures [20].

Accessibility and CHI

Kline and Glinert's *UnWindows V1* magnified areas of the screen to keep track of mouse position as a way to improve GUI navigation for users with low vision [15]. Obrenovic, Abascal and Starcevic [23] discuss accessibility in terms of multimodal design to propose a unified conceptual framework for accessibility in HCI. Kane et al [13] explore touch screen gestures, revealing that visually impaired users have different preferences for gestures than sighted users, preferring, for instance, gestures that used the edge of the screen as a reference. Sears and Hanson [29] offer a critical view of accessibility in HCI, pointing towards the wealth of studies that for a multitude of reasons lack representative users.

Haptics

Haptics for Accessibility

Sjöström proposes guidelines for Haptic Interaction Design for users with visual impairments to navigate GUIs [31]. Park and Howard [27] integrate sound and haptics to allow visually impaired users to explore and interact with remote environments immersively via a telepresence robot. The *Handscope* [14] is a haptic device that allows the visually impaired to experience statistical graphics on websites. Manshad et al focus on assisting distance learning for the visually impaired using a set of tangible interfaces [19]. Doush et al [2] describe a multimodal system enabling access to spreadsheet software for the visually impaired, combining an off the shelf 3D haptics device (the Novint Falcon), speech cues, and audio.

Talking about Haptics

Obrist [24] discusses the difficulty users have in describing haptic experiences. Atkinson et al discuss the difficulties in articulating tactile feel through interactive media [3].

Haptics in Music

The Moose [25], was "a haptic device aimed at providing access for blind sound engineers to the graphics-based computer interfaces currently found in digital sound studios." It functioned as a mouse enhanced with force feedback to allow the user to feel on-screen interface objects. The *D'Groove* is a DJ turntable that uses interactive force feedback on turntables [7]. It exploits the learned skills of a turntablist, presenting a variety of modes to allow the performer to "feel" the beat, or feel resistance in proportion to the amount of energy at a particular point in a record. Verplank's *The Plank* [37, 36], a haptic controller for music performance, has been adopted by the Copenhagen Institute of Interaction Design's (CIID) Music, Machines, Motors (M&M) project [5], making the technology available to the DIY community using modified Arduinos and was used in our project.

The work presented here extends these previous projects. We start with a general purpose interface like Doush, and creates a bespoke interface as the Handscope did. Whereas those works either translated elements of the user interface, or represented static graphical information, we seek to represent dynamic, time-based media in the haptic domain. This most closely matches the goals of The Moose. It, however, was based on the interaction paradigm of an existing input device (the computer mouse) adding force feedback to convey information about the graphical waveform. In our work, we directly translate sound to another sensory modality, bypassing GUIs and existing interfaces, to create a new interface that allows direct access to a haptic cross-modal mapping of audio content. By bypassing the visual domain, we avoid translating paradigms of the visual into the haptic modality.

METHOD

We conducted a series of three workshops over the course of the development of the idea and produced one technology mock-up and two prototypes. The final prototype of the Haptic Wave was deployed in a series of extended studio trials. The project has spanned a period of two and half years to date.

- Scoping Workshop: Meeting users, initial problem setting, brainstorming
- Workshop 2: Testing and gathering feedback from an initial "lo-fi" prototype
- Workshop 3: Testing of the final prototype responding to design ideas suggested in Workshop 2
- Studio Trials: Six participants lived with the Haptic Wave for up to 5 weeks, using it for everyday tasks in their studio, and reporting via a diary and interviews.

We worked with 11 users (9 men, 2 women; 8 U.K., 3 U.S.) through the series of activities. Some took part at the beginning, but were not available for later activities. Others were geographically distant and did not take part in all the group activities. From the eleven total, there were 3 who took part in all of the activities listed above.

Our users covered different areas of music and audio production. They were not all constrained to one working methodology, one audio practice or one musical genre. They ranged from soundtrack composers, session musicians (bass guitar, drummer), e-book and podcast producers, home studio amateurs, to music and performing arts teachers.

Phase 1: Scoping Workshop

We organized a one day scoping workshop and invited a group of eight of our users (7 male, 1 female) from the London area. The aims of this workshop were to: Meet the community we were working with; Find out their existing methods of working; Identify problems they experience in their work; Introduce new technologies in the form of mock ups; Brainstorm ideas for possible new solutions.

This structure was intended to relax the group, open up a dialogue about potential problems and then introduce them to novel technologies. It was then hoped that in the latter part of the workshop, the participants would be inspired by the technologies to suggest design ideas that we could work with.

Existing Work Methods and Needs Identified

Recording studios are very physical, tangible spaces, with a multitude of devices, and knob and slider controls. Our users take advantage of the physicality of space to overcome their visual impairments. However their descriptions of software use focused on touch-typing and creating scripts for shortcuts to standard editing commands.

Most of the users were accustomed to using off-the-shelf accessibility tools, such as screen readers, to aid computer use. Screen readers provide access to GUIs to visually impaired users by invoking voice synthesis to read out all menus and clickable buttons on a screen. Our users overwhelmingly felt that screen reader output was ill suited for audio editing, and resulted in unintuitive feedback about the sound they were trying to edit. This was compounded by the fact that the output of a screen reader uses the same modality (aural) as the content they were working with causing information about content to clash with the content itself.

The users preferred keyboard commands and shortcuts over using a mouse, which requires a readout of cursor position on the screen. While this was a suitable substitute for clickable buttons and menus (albeit laborious with many keystrokes required to do simple tasks) a keyboard could not substitute for a mouse in operations such as drawing curves or dragging to select ranges and regions.

Participants described existing use of motorized faders on commercial technology, such as the Behringer BCF2000 and the PreSonus FaderPort, to feel EQ curves. The participants are members of a community who are highly proactive in improving the conditions of their profession. Some have adopted DIY solutions such as an Raspberry Pi based screen reader for the open source Reaper audio editing suite[2].

Participants identified a broad range of issues they encountered in audio engineering tasks. Gaining a sense of the whole in the form of an overview arose as one difficulty. One user had difficulty managing multiple tracks, noting,

"Because I visualize the screen in my head, and then what I have in front of me doesn't reproduce this, what I'd like is something in front of me that actually reproduces in an accessible form what the screen looks like." (Sherie)

One workshop participant who worked with audio books brainstormed ways to scan through waveforms in order to find gaps, which normally represented chapter headings. They all expressed an avid interest in, what to them were bluesky, technologies such as touchscreens with tactile feedback.

Technology Mock Up

We used a general purpose haptic interface, the Novint Falcon[3] (Fig. 2), to create a demonstration of audio navigation aided by vibratory feel. We mapped the amplitude of the audio signal to the amount of force needed to move the Falcon, with louder parts providing more resistance. By moving the articulated arm of the Falcon interface from left to right, participants were able to navigate through a soundfile from beginning to end and "feel" its amplitude at any given point.

Figure 2. A workshop participant uses a general purpose haptic device to navigate audio.

Scoping Workshop: Results

Despite their experience with accessibility tools, users felt that off-the-shelf solutions were ill suited for sound editing, given the the highly graphical nature of audio software. The proliferation of parameters such as volume, panning, and multiple bands of frequency equalization mean that there is a myriad of interface elements for a screen reader to read out, taking an inordinate length of time. Moreover, the actual amplitude contour of the audio file being edited simply cannot be "read" as spoken text, leaving the main component of visual editing software untranslated for the visually impaired user.

The participants demonstrated an enthusiasm for haptic interaction. The notion of tactile feedback was present in their general consciousness, and they understood this as a way to feel parts of the interface. Ideas generated in the brainstorming included imagining a tactile array for a touch screen, similar to Bau et al [6]. By basing their imagination on their existing knowledge and experience, this typically meant imagining a tactile screen where buttons and clickable items could

[2]http://www.reaaccess.com

[3]http://www.novint.com/index.php/novintfalcon

be felt. This in effect extends the screen reader, but is limited to the notion of providing access to the GUI.

The Falcon demo enabled participants to use the sense of touch to directly access qualities of the sound. This triggered the imagination of participants, but the preliminary nature of the test raised as many questions as it answered. As a general-purpose haptic interface, the Falcon is not specifically designed to navigate time based media, where time typically is one dimension in a 2D representation. Instead, the Falcon is an articulated arm with six degrees of freedom, allowing free space motion. Our demo only used the horizontal axis to navigate time, with haptic feedback representing audio amplitude. Not only were the other dimensions of the device unexploited, the free space positioning left users without a point of reference, and caused undue tension as they tried to constrain their hand movement to the relevant axis. This had the negative effect of disappointing the participants' imagination of the potential benefit of haptic interaction with sound.

A discussion between three participants revealed a desire for haptics but a preference for 2D:

Peter: "You'd still want the force feedback, but in 2D rather than 3D."

(all): "Yep"

Scott: "Theres just too many degrees of freedom"

Steve: "It's that kind've 3D thing... my concept of that space, if I've got something to actually reference that against... that to me, it's much more tactile..."

Phase 2: Lo-fi Prototyping

The scoping workshop confirmed that our users were engaged, proactive, and already productive professionally despite frustrations with existing accessibility technologies. It pointed out the need for interface solutions that were not off-the-shelf and generic, but specifically designed to accommodate the sophisticated nature of their work and the specific characteristics of their medium – sound. The initial haptic mock up inspired some, but due to its lack of specific constraints, led to mixed reactions. Nonetheless, we observed a potential benefit to bypassing the graphic user interface to render the medium of sound directly in the haptic domain.

The Idea

To imagine a mapping of audio to the haptic domain for the purposes of editing sound, we looked at the qualities which made visual waveform editing compelling. A visual waveform is a Cartesian representation, with time in the X axis and amplitude in the Y axis. Silence is the middle of the Y axis, with the waveform increasing in amplitude, generally symmetrically above and below the origin (Fig. 1).

The width of the GUI displays a segment of time, with temporal resolution proportional to the zoom level. One could view the amplitude contour of an entire symphony, or zoom in to examine the envelope a single drum hit. In common audio software, clicking at a specific point in the timeline allows the user to select the point where an editing operation is to be applied. Dragging across a region from left to right generally allows a region to be selected to be cut, copied, or looped.

If the Falcon had too many degrees of freedom, we sought to reduce the dimensions of our proposed haptic design. Analog tape is one dimensional, with time running along the length of the tape, and audio untranslated to another modality. A visual waveform has two Cartesian dimensions. Could we design an interface with time represented on the horizontal axis and audio amplitude in another, haptic dimension?

Lo-fi Implementation

We used a disused scanner to produce a lo-fidelity prototype (Fig. 3). We dismantled it to use its carriage, base, and rails to provide us with a movable platform on which we mounted two motorized sliders (Y axis). An infrared sensor along the X axis was used to estimate the position of the carriage.

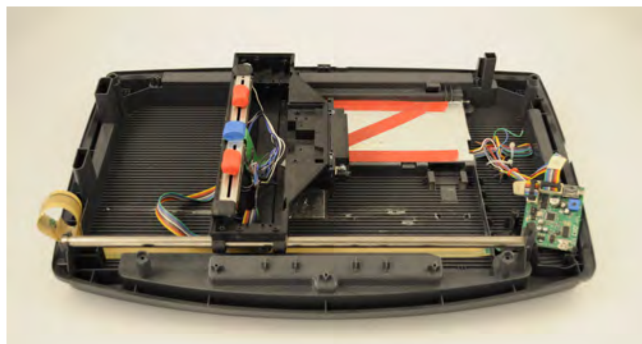

Figure 3. Lo-fi prototype using a repurposed scanner.

The two 6cm sliders mounted on the carriage replicated the positive and negative span of amplitude in the audio signal. The center point was marked with a knob, providing a point of reference for zero amplitude. The horizontal span of the device was constrained by the length of the rail of the scanner carriage. We worked with a hardware hacker to realize this lo-fi design, and interfaced it over USB with an Arduino sketch[4]. The USB serial data fed software written in the MaxMSP audio programming environment, which read a soundfile and allowed "scrubbing" (manually moving through the audio) using the position of the carriage[5]. The Arduino code moved the sliders to a point corresponding to the amplitude of the waveform at a given point in time.

Workshop 2: Testing the Lo-fi Prototype

Workshop 2 involved three participants who had taken part in the scoping workshop interacting with the lo-fi prototype in various ways (Fig. 4). We loaded in different audio samples and encouraged participants to scrub through the audio and feel the amplitude of the waveform. We designed a second scenario where the two sliders were used to convey two independent sources of information: waveform amplitude and mix automation. Participants also used the prototype to feel automation curves for panning and gain.

We discussed the participants' experience and impression of the device, such as how easy it was to gain information about

[4]https://www.arduino.cc

[5]https://cycling74.com/products/max/

sound through touch. An enthusiastic response came from a participant who is an editor of audio books, and therefore often looking for silences in long tracks, who noted,

"I NEED this. Can I take it home with me?...I'm editing an audio book for a client at the minute... We've got some chunky silences in there... and I'm having to do everything by listening." (Sherie)

Figure 4. One of our participants uses the lofi prototype in the workshop.

Another had mixed reactions, enthusiastic about the idea, but unfamiliar with haptics, and unsure about build quality:

"It's cool, but the fader action isn't very smooth, is it? ...It's quite hard to keep track of when you're moving... OK, yeah I get it. Yeah, it's cool, I dunno, I guess... It would be mega useful for kinda zipping through waveforms, tidying up beginnings and ends of stuff... There's just something about having your actually hands involved thats cool... I haven't really got the experience of taking things in by touch... I can imagine it becoming an instinctive way of working ... I can imagine having something vaguely like this on my desk and just reaching for it instinctively like how people grab for a mouse." (Scott)

A design point that emerged from the workshop was that two sliders (representing the top and bottom of the audio waveform) did not translate from the visual to the haptic. This was due to two reasons: 1) the pertinence of the source metaphor, and 2) ergonomics. One participant describes the two slider as,

"It's interfering, you're getting information from both the outside fingers... If it's just that [uses one slider] then I'm getting the simpler information." (Peter)

The two slider layout, with a reference point in the center, required the user to position three fingers in alignment along the vertical axis – the index finger on the middle reference, with the middle finger and thumb on the two sliders spanning the center, while simultaneously moving the hand to scan through the audio file. This required a twisting of the wrist that was unnatural for the user.

More importantly, one user indicated that he wasn't aware that a visual waveform was generally symmetric about the X axis:

"I didn't even know that the inverse below waveform existed!" (Scott)

While the "symmetrical" waveform has been adopted in most visual representations of audio, the fact that the our participant had never seen a graphical waveform representation means that this comparison was lost on him, despite his years of work as a professional audio engineer. Meanwhile it was intuitive that along the Y axis, a higher slider position meant greater amplitude. This demonstrated that a mapping of audio amplitude from the aural domain to the haptic domain did not benefit from comparison with a similar mapping to the visual domain. The two slider model generally made the device harder to use, and did not help convey more information to the participants.

THE HAPTIC WAVE PROTOTYPE

The feedback we received from users encouraged us to move forward to make a fully functioning prototype. Their reactions to the lo-fi prototype confirmed that a 2D representation was intuitive, but that a symmetrical representation around a center zero was unnecessary and ergonomically awkward. While different tasks which exploited the two slider configuration to display two different streams of information (such as waveform amplitude and mix level) alleviated the user of the uncomfortable single handed operation, the complexity was difficult for users to parse intuitively.

Figure 5. Building the Haptic Wave.

The Design Workbook

We wanted to formalize the information gained from workshop observations and user feedback before settling on a design specification for the final, high specification prototype. We worked with an industrial designer, and adopted the Design Workbook method [10]. The designer was new to the project, and was not a specialist in either audio or accessibility. This gave us the opportunity to have a fresh regard on the project, but posed the challenge to get him up to speed on the insight we had gained thus far. The workbook served as a kind of lab notebook documenting the design thinking process, and as a scrapbook for collecting images and ideas from an open ended generation of ideas. As (an electronic) written document, it served as a reference point for collaborators to focus, to communicate the project, and is archival. There was no planned structure for the workbook, and its resulting 56 page form emerged organically and consisted of sections on: *Participant insights; Design rationale; Related work; Textures and materials; Hardware development; Interactions; Form; Knobs and buttons*. The design workbook brought together background information, represented project stakeholders, and gathered ideas and sketches for the final prototype, and is reported in [28].

The Design

For the final prototype (Fig. 6) the designer, in discussion with the authors, settled on a rectangular form (50 x 19cm), compact enough to integrate into a music studio alongside other hardware, but wide enough for a comfortable left/right sweep with one hand, and the height of a single 100mm throw slider. The top was slightly canted towards the user. The base was fabricated in varnished birch wood (warm temperature, smooth wood grain). The slider carriage was surface mounted and oval shaped, affording grasping. It was 3D printed in ABS plastic (neutral temperature), and had a series of hatch marks along the side demarcating the upper and lower halves and indicating the vertical center point. The knob on the fader was round and indented with concentric rings, affording resting a finger, and fabricated from copper (cold).

Figure 6. The Haptic Wave final prototype with associated software.

The interior contains the following main components:

- 100mm Alps motorized slider mounted in the carriage
- Two parallel stainless steel rails, upon which the carriage is mounted
- Linear optical counter with a resolution of 3000 steps to measure the horizontal position of the slider carriage
- CIID M&M modified Arduino

Considerable calibration and a high degree of fabrication precision was necessary to allow smooth horizontal sweeping regardless of what part of the carriage the user grasped. In addition to the left-right scanning/scrubbing of audio, the user needs to mark specific points in the span. We explored different clicking mechanisms, making the knob a clickable button, or a simple conductive surface. Other possibilities included placing a micro-switch on the side of the carriage to allow clicking by grasping, or putting a button along the top surface of the device housing. The solutions that maintained single handed use were prone to accidental clicking, and other possibilities either required two handed use or required the user to let go of the carriage. In observing the workflow contexts in which users would integrate the device, their other hand was available, and a computer keyboard was omnipresent. All of the blind users with whom we worked were very comfortable finding specific keys on the keyboard by feel. Based on this, it was decided that instead of imposing a button solution on our device, that use of the device in conjunction with a keyboard or numeric keypad would allow more flexible reprogramming and multiple button inputs.

The Software

Software components consist of:

- Arduino code to report the horizontal position of the slider carriage from the optical counter, and to set the vertical level of the motorized slider according to audio amplitude, communicating via USB with the host computer
- Cross platform, standalone audio editing back end developed in MaxMSP to perform: load sample, mark start/end points, insert silence, and export edited audio
- Launcher that invokes the editor from standard DAW software[6] via a .bat file (PC) or Automator shell script (Apple)

As vertical position of the slider represents audio amplitude, the slider dynamics needed to be programmed and fine tuned to correspond convincingly to horizontal displacement of the carriage. A "force" parameter in the M&M code provided a sense of audio "pushing" its way up in amplitude. For consistent amplitude rendering, the audio is low-pass filtered to smooth slider movement and to fit the zoom level of total time to the width of the device. These dynamics need to be reconciled with static level representation so that the slider consistently achieves the same level at the same horizontal point, and attains that same point whether sweeping left or right.

The editing tasks that could be performed from the Haptic Wave were simplified from the lo-fi prototype, and focused entirely on waveform operations, with no secondary mix automation. Using the computer keyboard in conjunction with the device, users could mark edit in/out and loop start/end points. The core functionality became focused on operations where audio editing was typically facilitated by visual waveforms – finding silences, and identifying starts and ends of salient portions of a sound recording.

THE HAPTIC WAVE IN THE WILD

Evaluation of the Haptic Wave took part in two phases: A lab based workshop similar to the assessment of the lo-fi prototype; and technology probe [12] trials in our users's studios.

Workshop 3

The workshop (Fig. 7) involved three of the users giving feedback on the final prototype, and had a two-fold purpose: 1) To gauge users reactions to the device, and 2.) To discuss real-world usage scenarios to help us fine-tune the software backend for deployment in longer studio trials.

[6]A DAW is a Digital Audio Workstation, compromising hardware and software and a computer running Pro Tools, Logic, Reaper or similar software, forming the heart of the modern music studio.

The three users in this workshop had all previously tried the lo-fi prototype, over a year prior. One of those participants had the following reactions to the new prototype:

"It gives you this immediate, intuitive indication... to me, it must be pretty much as good as a sighted person would get looking at the waveform... It's great, I mean it's like a dream, it's brilliant. I really like it, because... when I'm editing normally, I have to do it entirely by ear... This would be so much quicker and more intuitive... The editing is superb... it's another way of representing haptically the level... If only there were a way to interface it simply [in the studio] I'd want it now, because of that quick way of finding the zero crossings and the lowest level and so on." (Peter)

Figure 7. Steve tests the Haptic Wave in the lab workshop.

Customizing the Haptic Wave

We settled on a setup where users would export a file from their DAW to be edited using the Haptic Wave, the resulting edit to be imported back into their production. The Haptic Wave allowed marking in/out points, cutting, pasting, and looping. Some customization, for example specific short key assignment or inserting silences, was left open to adapt to specific types of editing a user might want to perform.

We believed that a trial of the Haptic Wave was only useful if the user actually used the device as part of their professional and everyday work. Therefore, the Haptic Wave had to fit into their existing workflow. We adapted the software, including modifying keyboard shortcuts for specific users to be similar to their personal editing templates, for the device to be able to perform the tasks in the scenarios described above. The lab workshop allowed us to generate the mini-scenarios of contexts in which the Haptic Wave might operate, for instance situations where the Haptic Wave would be invoked as an editor from within a host DAW.

Studio Trials

We deployed the Haptic Wave in a series of 6 studio trials in London, Macon Georgia, Boston Massachusetts, and Ann Arbor Michigan (4 men, 2 women) in contexts ranging from home studios (Fig. 8) to professional recording facilities (Fig. 9). To the extent possible, a member of the team went to install, and later de-install the Haptic Wave. When he did so, he also documented the setup with photos and video, and conducted in-going and out-going interviews. In two of the cases (Boston, Ann Arbor), a face-to-face visit was not possible, and we aided the users to install and set up the device by email and telephone. Participants used the Haptic Wave on their own without supervision for a period of 1 to 5 weeks, integrating it into their daily work flow. They were asked to keep a diary, logging the times they used the Haptic Wave, the task they intended to do, and their experience in using it.

Figure 8. Steve uses the Haptic Wave in his home studio.

Results

In his diary, one user, Joey, describes the Haptic Wave in reference to sighted users:

"What the Haptic Wave represents is the ability for a blind person to use their hands like a sighted person would use their eyes." (Joey)

Another describes the effect of the device in terms of representations and goes on to discuss the mental image it evokes:

"I think it gives you a – it's not a visual representation, but I would call it a linking representation... This device...would give me that confidence to be able to know exactly where I am in the waveform and what the waveform looks like, it's giving me a picture in my head." (Steve)

Users were able to integrate the Haptic Wave into their work flows and commented on the speed and accuracy it brought to their work:

"I was able to do what I wanted to do, with the file, I was able to take out all the breaths and stuff like that... It worked, it did exactly what I wanted... editing this vocal track in regard to getting rid of extraneous noise has worked like a charm...It was very easy to find the start point, very easy to find the stop point, very quick, very easy, I grabbed the sample I needed, exported that, and then brought it in up over the sound effects and it was awesome!" (Joey)

One user saw the Haptic Wave as an effective way of speeding up editing, noting:

"It will definitely speed up the audio editing process quite a bit. For example, when cleaning up tracks, people like to remove unwanted noise like excessive breath or bleed from other sources. This is [a] very repetitive process, but [the] Haptic Wave will simplify it and shorten the time." (Chi K.)

Another user similarly noted that Haptic Wave sped things up, essential for his work:

"It's actually quicker [than existing techniques], because you've got much more tactility, it's just like using the scrub wheel... but it's much more accurate because it's actually giving you the feedback of the amplitude of the waveform as well... You can get much more control with a device like this because you've got much more information going, so you can get much more accurate feel where you are in the waveform... When I've actually done fine editing of snare drums and kick drums and things like that which I've done a lot of editing recently, its been amazing the results I've had... it's been really interesting to see the results I've had as opposed to when I'm not using it, its not that I can't do it, but I think this has made a big difference to the accuracy and also the speed. You know in the studio I'm a non sighted person [but] I have to be able to work as fast as a sighted person, and that's just the reality of the industry I work in... I think other people are using it for more speech based stuff say in broadcast situations, they are going to use it mostly for editing out breaths between vocals, chopping out bits for radio interviews, for vox pops or whatever, that will be really quick for that, and I know that from my experience of editing radio interviews." (Steve)

Figure 9. Joey with the Haptic Wave in his professional studio.

DISCUSSION

The process of research, design, and development of the Haptic Wave led to insights at several levels: 1.) Insight about our users; 2.) Methodological insights about working with visually impaired users and the prototyping process; and 3.) Insights about cross-modal representations in haptic interaction. The Haptic Wave as an artefact was a focal point throughout the process, at times the object around which research thinking took place, at times our response as researchers to issues raised by participants in the workshops, ultimately becoming the embodiment of the knowledge we all gained from the project.

As a prototype, the Haptic Wave worked in enabling our users to overcome their varying abilities to see, and edit audio through the sense of touch. We have demonstrated this in a lab based trial followed by extended studio trials *in situ* and gathered user reactions through qualitative feedback. For the Haptic Wave, a quantitative study would have been inappropriate to gather the data we needed or serve the community we were working with. Each of our users had highly personal approaches to being productive in professional situations with their own sets of off-the-shelf and custom accessibility tools. There would have not have been a single point of comparison for a task-based performance study. Moreover, personal studio technique is developed over time for all users, sighted or non-sighted, some taking years to develop a combination of studio configuration, keyboard shortcuts, and software utilities. A new device such as the Haptic Wave, takes time to learn, making it difficult to compare directly with the highly personal techniques developed, sometimes over years, by users.

Furthermore, Oakley has shown that haptic interaction does not always improve task performance time, but has the benefits of reducing error [22]. Despite this, and the time needed to learn a new device, it is interesting to note how much our users discuss speed benefits in their diaries. Steve mentions accuracy and speed together, with speed of execution being a metric by which he is compared professionally with sighted audio engineers.

Meanwhile, taking a qualitative approach brings with it its own problems and risks. The novelty effect of trying a new, cool, technique could bias our users' responses. However, alongside a general enthusiasm, we noted at every prototyping stage a healthy skepticism amongst our users, with Scott the most skeptical. His initial reaction to the haptic prototype, and feeling the motorised sliders move as he scrubbed through the waveform, was:

"Wow, that is weird!"

He described operating the first two fader prototype as:

"An RSI nightmare waiting to happen."

This was a key moment in our decision to abandon the two fader layout in favor of the single fader.

Scott did have more positive remarks as the Haptic Wave developed, enjoying the accuracy of the scrubbing in the final prototype, and remarked,

"I think this whole moving thing... does give you a degree of accuracy that you might not get unless you had a particularly good scrubbing [mechanism]."

When the motorized slider revealed a "quiet" sounding part to actually be loud, he remarked, *"Is it useful? Yeah, definitely."* When we subsequently asked him whether he could imagine the Haptic Wave being useful for finding non-zero crossings and audio editing, his answer was *"Yeah, definitely."*

With the acceptance of the device amongst our users, and given its increased familiarity, future work could include a quantitative task-based study to evaluate editing completion time using the Haptic Wave compared to their existing working methods. This would supplement the rich qualitative data gathered in the present work to help refine the design for potential commercial production of the Haptic Wave.

User Centered Design for Visual Impairments

The creation of functioning mock-ups and physical, working prototypes throughout the research was methodologically important in two ways – first with regard to the specific abilities of our users, and second as part of the design process We sought to include our community of users in the research through the use of user-centered and participatory design techniques. We imagined, at the outset, using structured brainstorming and ideation activities we have successfully previously used in imagining interactions with sound [9] and interactive music devices [32]. It was at the first scoping workshop where we understood that the highly visual techniques of user-centered design needed to be adapted to work with partially sighted and blind users.

We had anticipated this to some extent by bringing dictaphones to serve as "audio Post-It notes." Oral discussion was indeed important, in the workshops, and later in the diary taking activities. Diaries were either voice memo recordings or text based. Building lo-fi objects, beyond foam core dummies, but with basic movement and electronic functionality, became part of our dialogue with the participants.

Material Thinking

Physical objects became an important way to replace objects such as drawings in ideation activities. By touching a device, the user could comment on it. By using rapid prototyping software interfaced to the devices, we could alter or fine tune user interaction on the spot to try new settings or configurations, quickly getting feedback. Fast turnaround within workshop sessions, became embedded in the subsequent development of the prototypes between sessions, and constituted an iterative method where the artefact is in continuous evolution and forms the basis of research dialogue. This is a form of material thinking where we converse through objects, and where the knowledge we gain is embodied in the object [16].

The development of three levels of prototype, from technology mock-up using off the shelf hardware, to lo-fi prototype using DIY components, to a final polished prototype, was crucial in adapting and refining cross-modal interaction to be effective in real-world professional audio production contexts. It put in place a developmental process where feedback from users could be integrated and combined with our own specialist knowledge as sound researchers, allowing us to incrementally confront and solve issues such as the number of degrees of freedom, the number of simultaneous functions displayed, and the ergonomics of hand position.

The end result is not a lab prototype that validates a particular interaction modality in a hypothetical situation, but a fully functional device that can sit alongside highly advanced equipment in a professional recording studio, to study the benefits of our proposed mode of interaction in a real world setting. Having this level of finish with the prototype permitted us to take it into a range of such settings, from home studios, to radio editing suites, to music production studios.

This had a knock-on effect beyond the immediate research that, while good for project publicity, had the potential negative consequence whereby the device was misconstrued as a commercial product prototype. A television news programme covered the deployment of the device in the pro studio in Macon. While this was rewarding and points out the potential societal impact of the research, it is not a validation of the research *per se*. During our workshop sessions, users began to give feedback as product development ideas. While this meant that the device was compelling enough to imagine as a commercial product, it had a slightly distracting effect in that users would start to concentrate on qualities such as form factor, and functionality in the form of product features, risking to take the focus away from discussion of haptic interaction.

Direct Manipulation Interaction

The Haptic Wave was successful in its original objective to translate the audition-vision mapping of standard graphical audio editing tools to a sensory, haptic mapping. While this draws on prior work in tactile interfaces, there are some fundamental differences. Haptic perception can be divided into two broad categories, cutaneous and kinaesthetic [17, 35]. Much of the work cited in the related works section apply cutaneous stimulation in a multimodal context to augment sensation, facilitate navigation of GUIs and other tasks. The Moose is kinaesthetic, offering gross force resistance for interaction with music, but as O'Modhrain notes, it is, "basically a powered mouse, giving the user the ability to feel the screen objects under the mouse cursor" [25].

The Haptic Wave is kinesthetic and provides a direct sensory mapping of audio to the haptic domain. It is a direct manipulation interface [11, 30] as an input/output device that displays audio content without mediation by semantic representation, allowing tangible interaction with that content.

In doing so, the Haptic Wave does not use haptics as a supporting sensory modality in a multimodal context, nor does it render an existing user interface (such as a GUI) accessible. Instead, it allows direct manipulation of audio, taking advantage of intermodal perception that can be at once inherent and learned [34]. This may explain the intuitive sense that users found along with the learning and acclimatization they experienced. In Workshop 2 Sherie notes, *"The more I do it, the more its making sense."* Describing precision editing tasks (drum editing being demanding for rhythmic accuracy), Steve notes, *"I've done some quite detailed editing with it on kick drums, snares, high hat replacement in parts, and found it very intuitive."* The Haptic Wave is a device that takes time to learn, yet is intuitive.

The User as Expert

Our users were at the center of the research, providing key insights. While we had worked in the context of social inclusion before [33], this was our first project working with users with visual impairments. We were able to attune ourselves to these users' specific abilities and needs through the participation in the project of researcher Tony Stockmann who is himself visually impaired. It was through this colleague that we understood prior work in the field and met the community of visually impaired audio engineers and producers. This introduction by a member of the community created a situation of trust with our users.

Nonetheless, there were illuminating moments as we became accustomed to working with the group. In recognition of, and out of respect for, their limited or non-existent ability to see, we initially avoided using visual metaphors in our spoken and written exchange with them. This made us reflect on the UCD activities mentioned above, and helped us realize how visual our descriptions of the world are. Contrary to our cautious approach, however, our users all engaged actively with the visual world. Inasmuch as one user mentioned was unaware that an audio waveform was symmetrical about the horizontal axis, he also described "forming an image" mentally. Other users described palpably the way sighted colleagues work in the studio. This showed that, despite sensory limitations, our users functioned in a broader visually driven society [1]. This realization helped us to understand the usefulness and even correctness of visual metaphors in discussions with visually impaired users, on subjects of form factor, materials, and physical layout. Whether or not our users were able to see, vision was a sensory modality that was an active part of the conversation as we explored translation of audition to haptic sensation. We hope that this insight will aid other interaction designers in working with visually impaired users.

Our users form a small but active international community who help each other, sharing information and resources. This helped extend the scope of our project from workshops with the local community to international deployment in our studio trials. This also meant that the users were well disposed to trying out new, experimental technologies. Meanwhile, we had to be mindful to temper their enthusiasm, navigating what were at times florid interview responses to focus on frank, research-relevant responses. One user became an evangelist for the project, blogging it, promoting it, and self-producing videos where he explained the Haptic Wave and its benefits in his own words.

Given the users' active participation in the project, they became more than "just users" who were passive subjects and beneficiaries of the research. Short of becoming researchers or designers themselves, they were nonetheless "actors" in a dynamic exchange of ideas where technology prototypes became the objects of dialogical exchange [28]. This reached a transformational end point in considering the future of the Haptic Wave. "Eyes-free" interfaces are beginning to emerge from research [8, 18] into recent commercial music products.[7] We conducted an informal questionnaire of sighted musicians' use of visual waveforms in performance and studio practice (53 respondents). While the vast majority rely upon visual waveforms, many indicated that the screen was a distraction to creativity. For example, as laptop DJing increasingly replaces vinyl records at gigs, the screen and associated waveform pose problems, including difficult visibility under stage lights. One respondent mentioned the Haptic Wave as a possible alternative to eliminate the "screen as barrier" between DJ and audience. This suggests that a haptic audio interface could be useful for sighted users. In this context, our visually impaired participants cease to be users and become experts who can, with their specific abilities, inform further development of the Haptic Wave for broader take up.

[7] https://www.ableton.com/en/push/

CONCLUSION

We have presented the design, development, and deployment the Haptic Wave, an input/output interface that renders audio data as kinesthetic information. The device was developed in an iterative, participatory manner with a community of audio engineers/producers with visual impairments. Our experience in working with these users, and work translating techniques of brainstorming and user centered design to non-visual domains will be of interest to interaction designers.

The Haptic Wave facilitates manipulation of digital audio by a cross-modal mapping from the auditory domain directly to the haptic domain. This extends other techniques that use multimodal tactile feedback to enhance the navigation, and translation, of graphic user interfaces.

This discussion of our project has relevance in different ways for different groups. For an audio software developer, we show that accessibility can entail the translation of audio content to sensory modalities other than sight, and that software and hardware can be designed to accommodate alternate, non-graphical displays and input/output devices. For an interaction designer, we show how the design process itself is haptic, and that visual methods from user-centered design can be adapted to facilitate visceral design exploration with visually impaired users and users of different sets of abilities. For the HCI researcher interested in multimodal interaction, we demonstrate that the translation of content directly to other sensory modalities, rather than through representation of other interfaces, can allow compelling forms of direct manipulation interaction.

The Haptic Wave was borne out in a series of increasingly sophisticated prototypes where dialog among researchers, designers, and users was facilitated by the artefact(s). The cycles of development, from off-the-shelf, to lo-fi DIY, to high-spec final prototype permitted us to remain open to design suggestions throughout the process, but assure successful integration in real-world professional work environments at the end. In this sense, the trajectory as described here represents the full spectrum of research, from identifying user need triggering an idea, through proof-of-concept, leading to fully functioning technology prototype and evaluation resulting in ultimate deployment in the wild.

ACKNOWLEDGEMENTS

We wish to thank the participants who took part in the workshops and studio trials, and our collaborators from Queen Mary University of London, Nick Bryan-Kinns, Oussama Metatla, and Tony Stockman. We thank Fiore Martin for software development, "Leafcutter" John Burton for the lo-fi prototype, Tom Mudd for the Falcon demo, and Martin Klang for hardware engineering. Goldsmiths colleagues Nuno Correia, Francisco Bernardo, and Alessandro Altavilla assisted with documentation of the studio trials. We would like to express our appreciation of a close collaboration with the designer David Cameron. This research is part of the Design Patterns for Inclusive Collaboration (DePIC) project, funded by the Engineering & Physical Sciences Research Council UK (EPSRC, EP/J018120/1).

REFERENCES

1. 2010. "The Courtesy Rules of Blindness". (2010). Retrieved September 21, 2015 from http://www.blind.net/general-information/the-courtesy-rules-of-blindness.html.

2. Iyad Abu Doush, Enrico Pontelli, Tran Cao Son, Dominic Simon, and Ou Ma. 2010. Multimodal Presentation of Two-Dimensional Charts: An Investigation Using Open Office XML and Microsoft Excel. *ACM Trans. Access. Comput.* 3, 2, Article 8 (Nov. 2010).

3. Douglas Atkinson, Pawel Orzechowski, Bruna Petreca, Nadia Bianchi-Berthouze, Penelope Watkins, Sharon Baurley, Stefano Padilla, and Mike Chantler. 2013. Tactile Perceptions of Digital Textiles: A Design Research Approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1669–1678.

4. Paul Bach-y Rita and Stephen W Kercel. 2003. Sensory substitution and the human–machine interface. *Trends in Cognitive Sciences* 7, 12 (2003), 541–546.

5. Jakob Bak, William Verplank, and David Gauthier. 2015. Motors, Music and Motion. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '15)*. ACM, New York, NY, USA, 367–374.

6. Olivier Bau, Ivan Poupyrev, Ali Israr, and Chris Harrison. 2010. TeslaTouch: Electrovibration for Touch Surfaces. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology (UIST '10)*. ACM, New York, NY, USA, 283–292.

7. Timothy Beamish, Karon Maclean, and Sidney Fels. 2004. Manipulating Music: Multimodal Interaction for DJs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04)*. ACM, New York, NY, USA, 327–334.

8. Stephen Brewster, Joanna Lumsden, Marek Bell, Malcolm Hall, and Stuart Tasker. 2003. Multimodal 'Eyes-free' Interaction Techniques for Wearable Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 473–480.

9. Baptiste Caramiaux, Alessandro Altavilla, Scott G. Pobiner, and Atau Tanaka. 2015. Form Follows Sound: Designing Interactions from Sonic Memories. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3943–3952.

10. William Gaver. 2011. Making Spaces: How Design Workbooks Work. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 1551–1560.

11. Edwin L Hutchins, James D Hollan, and Donald A Norman. 1985. Direct manipulation interfaces. *Human–Computer Interaction* 1, 4 (1985), 311–338.

12. Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B. Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, Nicolas Roussel, and Björn Eiderbäck. 2003. Technology Probes: Inspiring Design for and with Families. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 17–24.

13. Shaun K. Kane, Jacob O. Wobbrock, and Richard E. Ladner. 2011. Usable Gestures for Blind People: Understanding Preference and Performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 413–422.

14. Da-jung Kim and Youn-kyung Lim. 2011. Handscope: Enabling Blind People to Experience Statistical Graphics on Websites Through Haptics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 2039–2042.

15. Richard L. Kline and Ephraim P. Glinert. 1995. Improving GUI Accessibility for People with Low Vision. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '95)*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 114–121.

16. Carl Knappett. 2011. *Thinking through material culture: an interdisciplinary perspective*. University of Pennsylvania Press.

17. Susan J Lederman and Roberta L Klatzky. 2009. Haptic perception: A tutorial. *Attention, Perception, & Psychophysics* 71, 7 (2009), 1439–1459.

18. Kevin A. Li, Patrick Baudisch, and Ken Hinckley. 2008. Blindsight: Eyes-free Access to Mobile Phones. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 1389–1398.

19. Muhanad S. Manshad, Enrico Pontelli, and Shakir J. Manshad. 2013. Exploring Tangible Collaborative Distance Learning Environments for the Blind and Visually Impaired. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. ACM, New York, NY, USA, 55–60.

20. Oussama Metatla, Nick Bryan-Kinns, and Tony Stockman. 2008. Constructing Relational Diagrams in Audio: The Multiple Perspective Hierarchical Approach. In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility (Assets '08)*. ACM, New York, NY, USA, 97–104.

21. Sigrid Norris. 2004. *Analyzing multimodal interaction: A methodological framework*. Routledge.

22. Ian Oakley, Marilyn Rose McGee, Stephen Brewster, and Philip Gray. 2000. Putting the Feel in "Look and

Feel". In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '00)*. ACM, New York, NY, USA, 415–422.

23. Zeljko Obrenovic, Julio Abascal, and Dusan Starcevic. 2007. Universal Accessibility As a Multimodal Design Issue. *Commun. ACM* 50, 5 (May 2007), 83–88.

24. Marianna Obrist, Sue Ann Seah, and Sriram Subramanian. 2013. Talking About Tactile Experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1659–1668.

25. Sile O'Modhrain and Brent Gillespie. 1997. The Moose: A Haptic User Interface for Blind Persons. In *Proc. Third WWW6 Conference*.

26. Sharon Oviatt. 1999. Ten Myths of Multimodal Interaction. *Commun. ACM* 42, 11 (Nov. 1999), 74–81.

27. Chung Hyuk Park and Ayanna M. Howard. 2012. Real World Haptic Exploration for Telepresence of the Visually Impaired. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '12)*. ACM, New York, NY, USA, 65–72.

28. Adam Parkinson, David Cameron, and Atau Tanaka. 2015. Haptic Wave: Presenting the Multiple Voices, Artefacts and Materials of a Design Research Project. In *Proceedings of the 2nd Biennial Research Through Design Conference*.

29. Andrew Sears and Vicki L. Hanson. 2012. Representing Users in Accessibility Research. *ACM Trans. Access. Comput.* 4, 2, Article 7 (March 2012).

30. Ben Shneiderman. 1993. 1.1 Direct Manipulation: a Step Beyond Programming Languages. *Sparks of innovation in human-computer interaction* 17 (1993).

31. Calle Sjöström. 2002. *Non-visual haptic interaction design-guidelines and applications*. Lund University.

32. Atau Tanaka, Olivier Bau, and Wendy E. Mackay. 2013. The A20: Interactive Instrument Techniques for Sonic Design Exploration. In *Sonic Interaction Design*, Karmen Franinovic and Stefania Serafin (Eds.). MIT Press, 255–270.

33. Atau Tanaka, Lalya Gaye, and Ranald Richardson. 2010. Co-production and Co-creation: Creative Practice in Social Inclusion. In *Cultural Computing*, Ryohei Nakatsu; Naoko Tosa; Fazel Naghdy; Kok Wai Wong; Philippe Codognet (Ed.). IFIP Advances in Information and Communication Technology, Vol. 333. Springer, 169–178.

34. Esther Thelen and Linda B Smith. 1996. *A dynamic systems approach to the development of cognition and action*. MIT Press.

35. Jan BF Van Erp, Ki-Uk Kyung, Sebastian Kassner, Jim Carter, Stephen Brewster, Gerhard Weber, and Ian Andrew. 2010. Setting the standards for haptic and tactile interactions: ISOs work. In *Haptics: Generating and Perceiving Tangible Sensations*. Springer, 353–358.

36. Bill Verplank and Francesco Georg. 2011. Can Haptics Make New Music? Fader and Plank Demos. In *Proceedings of the 2011 International Conference on New Interfaces for Musical Expression (NIME '11)*. 539–540.

37. Bill Verplank, Michael Gurevich, and Max Mathews. 2002. The Plank: Designing a Simple Haptic Controller. In *Proceedings of the 2002 Conference on New Interfaces for Musical Expression (NIME '02)*. 1–4.