Spring 1-1-2016

Ambient Audio and Calm Technology

William Christopher Payne
University of Colorado at Boulder, william.payne@colorado.edu

Follow this and additional works at: https://scholar.colorado.edu/csci_gradetds
Part of the Computer-Aided Engineering and Design Commons, Computer Sciences Commons, and the Music Commons

Recommended Citation

This Thesis is brought to you for free and open access by Computer Science at CU Scholar. It has been accepted for inclusion in Computer Science Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
Ambient Audio and Calm Technology
Replacing Traditional Notification Systems with Blended Sonifications

William Christopher Payne

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirement for the degree of
Master of Science
Department of Computer Science
2016
This thesis entitled:
Ambient Audio and Calm Technology
written by William Christopher Payne
has been approved for the Department of Computer Science

____________________________
Professor Shaun Kane

____________________________
Professor Clayton Lewis

____________________________
Professor Michael Theodore

Date: November 17, 2016

The final copy of this thesis has been examined by the signatories, and we
find that both the content and the form meet acceptable presentation standards
of scholarly work in the above mentioned discipline.

IRB Protocol 16-0588
As technology becomes increasingly ubiquitous and mobile devices are joined and even outnumbered by wearables, the need for calm technologies, those that support and inform without demanding attention, grows more urgent. To help understand how audio may play a role in the development of such technologies, we examine a representation technique known as blended sonification which utilizes sounds already present in the environment. We describe the implementation of two prototype calm technologies, Sonification Station and Audio Poke, and report the results of two experiments in which music was used as a medium for blended sonification. Participants performed similarly on a transcription task when a traditional alert sound was replaced by blended music sonification as a form of notification, and participants across levels of music training were generally proficient at detecting and identifying audio effects applied to music.
<table>
<thead>
<tr>
<th>SECTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Overview</td>
<td>3</td>
</tr>
<tr>
<td>II. RELATED WORK</td>
<td>3</td>
</tr>
<tr>
<td>Sonification</td>
<td>3</td>
</tr>
<tr>
<td>Notifications and Productivity</td>
<td>5</td>
</tr>
<tr>
<td>III. PROTOTYPES</td>
<td>6</td>
</tr>
<tr>
<td>Sonification Station</td>
<td>7</td>
</tr>
<tr>
<td>Overview</td>
<td>8</td>
</tr>
<tr>
<td>Sonification Station Data Inputs</td>
<td>8</td>
</tr>
<tr>
<td>Sonification Station Sound Outputs</td>
<td>9</td>
</tr>
<tr>
<td>Evaluation</td>
<td>10</td>
</tr>
<tr>
<td>Audio Poke</td>
<td>11</td>
</tr>
<tr>
<td>Overview</td>
<td>12</td>
</tr>
<tr>
<td>Evaluation</td>
<td>13</td>
</tr>
<tr>
<td>IV. EXPERIMENTAL DESIGN</td>
<td>14</td>
</tr>
<tr>
<td>Experiment 1: Task Productivity</td>
<td>15</td>
</tr>
<tr>
<td>Design</td>
<td>15</td>
</tr>
<tr>
<td>Hypothesis</td>
<td>15</td>
</tr>
<tr>
<td>Implementation</td>
<td>16</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>19</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>20</td>
</tr>
<tr>
<td>Design</td>
<td>21</td>
</tr>
<tr>
<td>Hypothesis</td>
<td>21</td>
</tr>
<tr>
<td>Implementation</td>
<td>21</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>24</td>
</tr>
<tr>
<td>V. METHODS</td>
<td>26</td>
</tr>
<tr>
<td>Participants</td>
<td>27</td>
</tr>
<tr>
<td>VI. RESULTS</td>
<td>27</td>
</tr>
<tr>
<td>Experiment 1</td>
<td>28</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>35</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>VII. DISCUSSION</td>
<td>41</td>
</tr>
<tr>
<td>Experiment 1</td>
<td>41</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>43</td>
</tr>
<tr>
<td>VIII. CONCLUSION</td>
<td>45</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>46</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>52</td>
</tr>
<tr>
<td>Experiment 1 Pre-Survey</td>
<td>53</td>
</tr>
<tr>
<td>Experiment 1 Post-Survey</td>
<td>56</td>
</tr>
<tr>
<td>Experiment 2 Form</td>
<td>62</td>
</tr>
</tbody>
</table>
### TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Audio Samples Used in Experiment 2</td>
<td>23</td>
</tr>
<tr>
<td>2. Automatic Data Collection for Experiment 2</td>
<td>25</td>
</tr>
<tr>
<td>3. Crosstab Table Showing Participant’s Progress by Alert Type</td>
<td>31</td>
</tr>
<tr>
<td>4. Likert Scale Responses for Each Alert</td>
<td>34</td>
</tr>
<tr>
<td>5. Likert Scale Responses and Identification Data Taken from Experiment 2 Survey</td>
<td>39</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.</td>
<td>User interface for Sonification Station prototype app</td>
</tr>
<tr>
<td>2.</td>
<td>User interface for Audio Poke toolbar application</td>
</tr>
<tr>
<td>3.</td>
<td>Popup window displaying relevant browser history</td>
</tr>
<tr>
<td>4.</td>
<td>Primary task with participant transcription on left and original text on right</td>
</tr>
<tr>
<td>5.</td>
<td>State diagram for secondary task as used by participants</td>
</tr>
<tr>
<td>6.</td>
<td>Max MSP interface used by researchers</td>
</tr>
<tr>
<td>7.</td>
<td>Log file of experiment taken during the gradual alert condition</td>
</tr>
<tr>
<td>8.</td>
<td>Experiment 2 interface used by participant while audio sample is playing</td>
</tr>
<tr>
<td>9.</td>
<td>Audio samples rendered in Logic Pro X</td>
</tr>
<tr>
<td>10.</td>
<td>Estimated timeline of the experiment</td>
</tr>
<tr>
<td>11.</td>
<td>Bar graph showing type of alert against number of times the participant engaged the secondary task</td>
</tr>
<tr>
<td>12.</td>
<td>Bar graph showing alert type against the seconds it took participants to complete each secondary task</td>
</tr>
<tr>
<td>13.</td>
<td>Likert scale responses following Experiment 1 for each condition</td>
</tr>
<tr>
<td>14.</td>
<td>Errors participants made in Experiment 2 grouped by musical training</td>
</tr>
</tbody>
</table>
15. Number of correct identifications participants made in Experiment 2, grouped by musical training ........................................... 37
16. Participant detection times for audio samples when an audio effect was added or removed ........................................... 38
17. Likert scale ratings for Noticeability and Annoyingness Grouped by Audio Effect ........................................... 41
I. INTRODUCTION

In 1995, Mark Weiser and John Seely Brown coined the term “calm technology” describing it as simultaneously calming and informative, a “key challenge in technology design for the next decade” (Weiser & Brown, 1995). It is clear over two decades later that the need for calm technologies still exists, and if anything, the burden technology places on our daily lives is exacerbated. A 2013 study estimates people check their phones 150 times per day or about every six minutes (Ahonen & Moore, 2013), and a common hallucination dubbed phantom vibration symptom (Deb, 2014) demonstrates that technology may demand our attention even without producing an actual notification. Furthermore, a large body of software is meant to distract us, to make us rapidly scroll and obsess over the status of our applications because our attention produces income for software developers and advertisers (Bosker, 2016). The experience of most users is a far cry from Weiser and Brown’s conception of technologies designed to be ignored most of the time.

Numerous artifacts that do not rely on software provide inspiration for beginning to develop calm technologies. Inner office windows connect those inside the window with those outside via subtle cues, and do not mandate the interactions present at close proximity in an open workplace (Weiser & Brown, 1995). A tea kettle is unused most of the time and can be ignored, but as the water boils, noise gradually fills the room culminating with a whistle or a pop when it is ready. Weiser and Brown argue that calm technology engages both the center and periphery of attention, moving gracefully between the two (Weiser & Brown, 1995). Such technologies manage to relax us in two ways. They reduce the sheer number of tasks we have to actively consider and they automatically center the most critical task so it does not build anxiety in the background. A substantial body of work inspired by Weiser and Brown is the development of *ambient displays*, interactive sculptures that sit in the background of a person’s environment and convey critical information to be interpreted by a mere casual glance. For example, the CareNet Display is a picture-frame sized screen that summarizes the current needs and behaviors of an elderly patient for local members of the elder’s home care network (Consolvo, Roessler, & Shelton, 2004). Color coded icons give an overall picture (not always successfully) of the most critical patient data as ranked by
anxiety

In Background concentration, listening Ozawa, is task software software file phones technology. money is being spent. bracelet is photo safekeeping bracelet weekly and application specific information. caretakers, and the CareNet Display includes an interaction mode for caretakers to explore more specific information.

Other modern software examples of calm technology include f.lux (Herf & Herf, 2009), an application that changes a display's color to match the time-of-day resulting in reduced eye strain and more healthy sleep habits, and UbiFit, a mobile phone prototype that encourages users to reach weekly fitness goals through growing digital flowers in a garden scene wallpaper when they succeed (Consolvo et al., 2008). A wearable calm technology is the Disney Magic Band, an unassuming bracelet containing an RFID chip and radio. It does not need to be charged or placed into safekeeping before going on a rollercoaster, and it tracks park experiences such as the location of a photo taken or a broken ride which will automatically grant reimbursement. In the few occasions it is used, e.g. to get into the park or purchase a snack, the experience is pleasant, even magical as the bracelet sets a terminal to life with swirling, rainbow LEDs (Kuang, 2015). One almost forgets real money is being spent.

Music played in the background of a work environment is often used as a form of calm technology. It certainly is ubiquitous with 61% of millennials claiming to listen to music from their phones while 74% own an iPod or mp3 player (Zickuhr, 2011). The advent of mp3 players, digital file formats, and cloud streaming services has made private music listening possible in any setting, and studies show that it can increase productivity and positive mood in the workplace. In the software development community, new programmers with access to the most music performed the highest on a programming task, and data collected from the work environments of multiple software companies found that state positive affect and quality-of-work were lowest while time-on task was highest when music was removed (Lesiuk, 2005). Even the great author Haruki Murakami is known to work almost exclusively with a record on from his substantial collection (Murakami & Ozawa, 2016). A large scale UK study synthesizing work of the past 100 years and tracking the listening habits of nearly 300 office employees found that music can positively benefit inspiration, concentration, positive distraction, stress relief, and management of personal space (Haake, 2011). Background music simultaneously helps workers engage in the task-at-hand and provides an escape. In other words, like a calm technology it can remain in the periphery until the primary task causes anxiety bringing it to the center of attention. In today's open-office environments, background
music even presents a social cue to nearby workers who are not listening to the music. Similarly to blinds pulled down on inner-office windows, the act of wearing headphones signals a desire to be left undisturbed. Of course, music is not without limitations and context is key: some respondents found that music can increase distraction especially during work requiring a lot of concentration. Other respondents were nervous that listening could disturb their neighbors or cause them to look unprofessional if a client walks through the office (Haake, 2011).

Overview

Looking forward, in Section II we describe other influential work including an audio interaction technique called sonification and experimental systems that attempt to deliver notifications in a manner resulting in reduced levels of distraction. In Section III, we describe and evaluate two prototype software applications that utilize sonification to present information to a user. In Sections IV and V, we describe the design and implementation of two experiments in which participants used blended sonification of music to deliver notifications and were asked to identify audio effects that were applied to music. Finally, in Sections VI, VII, and VIII we discuss the results of each experiment and identify directions in which to move forward after this research.

II. RELATED WORK

Sonification

A common technique for representing information, and thus a potential approach for building calm technology, is sonification. Sonification may be defined as the “core component of an auditory display; the technique of rendering sound in response to data and interactions” (Hermann, Hunt, & Neuhoff, 2011). Numerous audio parameters are at the disposal of developers for creating non-speech sonifications including the sound’s frequency, loudness, timbre or quality of the sound, duration, and direction or position of the sound. Often, these parameters are manipulated via MIDI, a common protocol used to represent musical data (Billey, 2016). The audio itself may be produced in a variety of ways from wavetable synthesis, the simulation of natural sounds using existing audio recordings, to frequency modulation, the use of one signal as a modulator for another, to additive
synthesis, the stacking of very simple waveforms to produce more complicated sounds (Brewster, 2003).

Sonifications are useful in many contexts including those when a person is blind and cannot see a display, a person’s vision is otherwise occupied, or sound happens to be the best representation among other feasible approaches. Examples of auditory displays used in assistive technology include the widespread use of screen readers, tools that convert displayed text to synthesized speech and sometimes support nonspeech sounds, “Follow That Sound,” an experiment that found a combination of manipulating pitch and panning was effective in teaching touchscreen gestures to blind people (Oh, Uran, Kane, & Leah, 2013), and work at Microsoft: The Seeing AI project combines computer vision and natural language processing to describe a person’s surroundings (Linn, 2016), and Cities Unlocked aims to make urban areas more accessible for people with sight loss (Wainwright, 2014). Other examples of sonification include a feature in some cars in which the rate of a beeping sound symbolizes the proximity of objects in the car’s rear, the Sonification Sandbox, a toolkit developed at Georgia Tech that allows users to map data to multiple auditory parameters (Walker, 2009), and the music of Holly Herndon who used the sounds of internet browsing as compositional material (Đuverović & Lina, 2016).

More recently, researchers have proposed a framework called blended sonification, a form of auditory data representation that blends into the user’s environment without ever directly confronting him/her. A working definition is “the process of manipulating physical interaction sounds or environmental sounds in such a way that the resulting sound signal carries additional information of interest while the formed auditory gestalt is still perceived as coherent auditory event” (Tünnermann, Hammerschmidt, & Hermann, 2013). For example, the noise produced from water flowing from a showerhead may be filtered to indicate the time, or a doorknock may be augmented with a reverb effect symbolizing to the knocker whether someone is present or the length of time he/she has been absent. Blended sonification is a promising tool for designing calm technology since it takes advantage of existing environmental sounds that naturally inhabit a person’s periphery. An early prototype, the Augmented Keyboard, tracks the weather and applies a filter effect based on parameters like temperature and humidity to the sounds made by a user typing on a keyboard (Tünnermann et al., 2013). The prototype requires a contact microphone and
speakers to blend the filtered sound with the original. Another prototype *Upstairs* is intended to tighten the bond between partners in a long distance relationship: Contact microphones are placed on the floor of each person’s apartment while audio is streamed to surface transducers affixed to each ceiling simulating the effect of the individuals mutually living upstairs from one another (Bovermann, Rene, Leichsenring, & Hermann, 2012). The system provides peripheral awareness of the behaviors and rituals of each person. Finally, a shoe-based prototype intended to encourage a more active lifestyle alters the frequency spectra of the sounds a person produces while walking (Tajadura-Jiménez et al., 2015). The resulting sounds are consistent with those produced by a lighter or heavier body. In particular, augmenting the high frequencies in accordance with a lighter body can motivate more physical exercise.

One limitation of existing blended sonifications is that a technological setup, often requiring additional microphones, speakers, or sensors, may be burdensome for many users. Our work is meant to take advantage of the ubiquity of background music and produce a system that requires no external hardware fitting into many people’s existing music listening habits. Furthermore, blended sonifications have only been applied to simple, natural sounds which, contrary to produced music, do not already have audio effects applied. It is yet to be known if a blended music sonification is detectable, useful, or pleasant.

**Notifications and Productivity**

The experiment described later in this work explores music as a medium for blended sonification and applies the resulting sound to replace traditional audio notifications like ringtones. The approach is motivated by a well-documented need to reduce the cognitive load placed on workers by distracting notifications. For example, experiments by Cutrell, Czerwinski, and Horvitz have shown that instant messages are harmful to overall task performance such as using a menu, writing text, or evaluating search results especially if the notification is delivered when the user is beginning the task (Cutrell, Czerwinski, & Horvitz, 2001). Instant message notifications affect some tasks more than others. For example, a notification severely disrupts the task of searching a list for a particular item (Cutrell et al., 2001). Additionally, Speier et al found that interruptions lower performance of complex decision-making tasks, and the effect is worsened with multiple
interruptions and when interruptions have little to do with the task at hand (Speier, Cheri, Valacich, & Iris, 1999). Simply ceasing the use of notifications is not the solution. A recent field study ran at Microsoft Research found that some users who turn notifications off self-interrupt more to explicitly monitor email arrival, and users perceive greater value in the awareness of information that notifications deliver than the disruption they cause (Iqbal & Horvitz, 2010).

Research attempting to rectify the problem of disruptive, yet valued notifications has focused towards the development of context-aware systems which try to identify the individual user’s state and hold off on signaling distracting notifications until the right moment. For example, researchers at Carnegie Mellon ran a series of experiments that show basic sensors placed within an office environment can estimate human interruptibility as well as humans can who are watching a video of workers in the environment (Fogarty et al., 2005). Researchers at the University of Illinois developed a system called OASIS which uses machine learning to predict breakpoints, intervals between tasks, in a person’s workflow and deliver all current notifications during these breakpoints (Iqbal & Bailey, 2008). They found that users reacted more quickly and expressed less frustration when notifications were delivered at breakpoints rather than immediately.

While many experiments have looked at context-aware systems, intelligent ways to delay notifications, and support for recovering from disruptions, no prior studies have explored adapting the form of audio cue used for the notification. Ideally, an ambient display, with more subtle methods of delivering notifications, could allow the user to act on new information at natural stopping points without using any context detection. Keeping in line with Weiser’s principles for ubiquitous computing and calm technology (Weiser, 1991), this research explores the use of ambient audio that fades into the background and is detected once the user perceives it, in other words informing, but not demanding attention.

III. PROTOTYPES

We developed two prototype computer applications during the course of this work. The first, Sonification Station, is a notification system that communicates a variety of information to the
user with ambient sound. The second, Audio Poke, is a toolbar application that uses blended sonification of music to discourage unproductive online behavior.

First, in the following subsections, we describe the characteristics and implementation of each prototype. Second, we evaluate each prototype in regards to three criteria identifying a calm technology (Weiser & Brown, 1995):

1. **Movement**: A calm technology should easily move from the center of attention to the periphery and back.

2. **Reach**: A calm technology should enhance *peripheral reach* by bringing more details into the background.

3. **Temporality**: A calm technology should relay some sense of familiarity to a user and tune the user's periphery connecting him/her to the past, present, and future and to the world around him/her.

In addition to the descriptions below, video demos of the prototypes may be found on the researcher’s website, http://williepayne.com/projects/masters-thesis.

**Sonification Station**

![Sonification Station](image)

Figure 1: User interface for Sonification Station prototype app
Overview

Sonification Station is a Mac application designed to notify a user of external information entirely via audio stream (Figure 1). Sonification Station consists of input modules that monitor data sources and sound-producing output modules. Any available data input may be mapped to any sound output. Sound outputs may only be triggered by one data source, but a single data source may be represented by multiple sounds. For example, in one arrangement a user may map the temperature to the pitch of his/her key press sounds, the number of correctly spelled words he/she has typed in a work session to both wind and rain sound effects, and the amount of time spent on Facebook to a filter effect applied to music. Sonification Station was built using Max MSP (Puckette, 1991) for the interface and audio processing and Python for gathering data in real time.

Sonification Station Data Inputs

Included data sources help the user track the following behaviors and information:

● Productivity
  ○ Facebook use: If the user has an active internet browser window containing the domain, “https://www.facebook.com” a bar will fill in ten minutes. A full bar takes 30 minutes to clear, but if the user gets off Facebook, any affected sounds will clear immediately. Other distracting websites are not trackable, but additional settings could make this feature customizable.
  ○ Current typing rate: The rate of the typing is calculated by tracking the number of characters a user types per second.
  ○ Correctly spelled word count: Each time a user types a correctly spelled word, an overall count is incremented. Words may be typed in any application regardless of whether the Sonification Station window is active.

● Activities
  ○ Number of unread emails: If the user provides basic access to a Google Email Account, Sonification Station can track the number of unread emails in the inbox.
  ○ Time until next calendar event: Similarly, if access is provided for a Google Calendar Account, Sonification Station can track the number of minutes until the user’s next
scheduled event. This input, along with the following input, may help to inform the user of an approaching deadline using a stimulus that gradually becomes more noticeable.

- Time until one should leave to catch a bus: Taking advantage of the Google Maps API, if the user provides a starting and ending location, directions using public transportation will be gathered. The first direction, minutes until one should leave to walk to the first mode of transportation, is tracked.

- Other external information
  - Weather: When provided a zip code, the current temperature will be returned from an online source.
  - Percent CPU Utilization: The user may gain an idea for how demanding his/her applications are on the computer's CPU at any time.

Sonification Station Sound Outputs

Sonification Station represents the data listed above with multiple audio outputs that produce sounds in real time and gradually change in response to the changes in data. Generally, higher data values, e.g. lots of unread emails, produce louder, more discernible sounds. The following sounds have been implemented:

- Sonifications
  - Wind Generator: The sound of the wind is emulated with pink noise filtered by a dynamic bypass filter. Optional wind chimes will sound at a rate determined by the magnitude of the wind.
  - Rain Generator: Rather than just increasing the amplitude of a set rain sound, the rain generator shows the range of a burgeoning storm beginning with a few quiet drops and ending with a torrential downpour. The effect is implemented as a series of pulses that excite a range of noise band frequencies to emulate the sounds of varying droplet sizes landing on different surfaces (Farnell, 2010, p. 441).

- Blended Typing Sonification: Sonification Station can use the sound produced by typed keypresses captured from a microphone similarly to the Augmented Keyboard
(Tünnermann et al., 2013). Without an external microphone, key press sounds may synthesized when the user presses a key from preset samples including a typewriter and the tapping of glass. The sound of these keys may then be outputted with an additional pitch shift effect or reverb effect. If a sample is used, the audio produced is panned to match the position on a QWERTY keyboard that a key is typed.

- Blended Music Sonification: A low pass filter is applied to music playing through built-in audio player. A low pass filter effect is also used in the experiments discussed in sections IV-VIII.

All data sources are mapped to the same range before being outputted as sound. For example, ten unread emails, ten minutes on Facebook, and 0 minutes left until one should leave for the bus will all produce the maximum amplitude or effect level for a sound output. While these values were selected arbitrarily or in accordance with the developer’s personal preferences, further customization options could provide a user explicit control on how data is presented. Furthermore, all data is linearly mapped into a sound output, but other mappings, exponential or logarithmic, may be useful in matching input data with a desired output. For example, an exponential mapping of CPU usage could output imperceptibly low levels of sound for most values, but become quite loud if the CPU is working much harder than normal past a particular cutoff value.

Evaluation In Regards to Weiser and Brown’s Three Criteria Identifying Calm Technology

1. Movement: The technology moves between the user’s center of attention and periphery as a result of the sounds selected. Each is intended to be calming, subtle, and undetectable at low intensities unless the user explicitly turns his/her attention to the sound environment. One limitation is the implementation of Sonification Station itself. It is rather unwieldy as it must always be open to run and at best can be hidden or minimized. In other words, it is not capable of moving into the periphery as say F.lux which, after an initial set up, starts immediately when the computer is turned on and is ignored as a tiny icon until the user needs it (Herf & Herf, 2009). A better implementation would be as a widget or toolbar app that can be completely hidden most of the time after its settings have been specified.
2. **Reach:** Information is brought into the user’s periphery by converting data that the user typically would search for visually (e.g. opening an internet tab, looking at a phone) into information that is always present in the form of sound or lack thereof.

3. **Temporality:** The technology is intended to feel familiar as a result of using sounds produced in nature, by the user directly, or by the user’s music. The use of blended sonification connects the user to the world around him/her, and the use of gradual auditory changes occurring over time helps position the user within the past/present/future. For example, instead of counting the tones a phone has produced in a given window of time, a user may detect at time “x” that some sounds are are louder or more present than an earlier time “z” making possible a comparison between changing states of data. More extensible customization of how these sounds shift in response to dynamic data flow could support users even further.

**Audio Poke**

![Figure 2: User interface for Audio Poke toolbar application](image)

**Overview**

Following Sonification Station, we intended to create an application that was targeted to common needs and adapted from a tool already in use. The result is Audio Poke (Figure 2), an internet productivity application inspired by existing services, such as the open source SelfControl (Stigler & Lambert, 2016), that allow users to block distracting websites during specific times of the day. Audio Poke works similarly, but rather than preventing access, it applies an audio effect to the user’s system audio output in increasing severity as long as the user remains on the site. By using blended sonification of music instead of other notifications, Audio Poke is meant to promote healthy online behavior without contributing to notification overload, as detailed in the description.
It was implemented for the Mac Toolbar using Python for interface and website monitoring and Max MSP to produce the audio effects. Unlike Sonification Station, Audio Poke supports all system audio including music played from iTunes, Spotify, etc. This is accomplished using Soundflower ("Soundflower," 2016), an open source extension that provides support for routing audio between applications.

Audio Poke consists mainly of two components: browser monitoring/logging, and audio manipulation. First, the app enables a user to specify a list of unique time-wasting websites. Any time it detects an open tab with a domain from the user's list, it logs the exact time the tab was opened and similarly logs the exact time the tab is closed or browsed away from. Logs include only the domain and not specific pages, and websites absent from the list are completed ignored by the application. A full log of browsing information is stored online in a database implemented with Firebase ("Firebase," 2016), and the user may open a window displaying their recent history at any time (Figure 3). Web monitoring supports both the Safari and Chrome browsers.

![Figure 3: Popup window displaying relevant browser history](image)

Second, Audio Poke alters the user's audio with one of three selectable effects. As the user navigates to and remains on a time-wasting website, the magnitude of an effect will increase reaching a maximum value after ten minutes. The level of effect decreases at a slower rate taking thirty minutes to vanish completely from a maximum state. However, to reduce frustration, the user will only hear an effect if he/she is currently on a blacklisted website. For example, if the user spends ten minutes on Facebook, an unproductive site, the effect will increase until it is applied in
full force. The instant the user navigates to a productive website, the effect will vanish within seconds, but if after a few seconds the user navigates right back to Facebook, the effect will jump right back to where it left off. The implemented effects include:

1. **Low-Pass Filter:** The cutoff value is lowered from an imperceptibly high frequency to a very noticeable frequency, about 150 Hz. In effect, the user’s audio changes from virtually unaltered to slightly muffled to almost inaudible.

2. **Bitcrusher:** A signal quality reducer lowers the sampling rate of the audio input from full to nearly zero (“degrade~ Reference,” 2014). As the effect is magnified, the audio ranges from unchanged to slightly warped, e.g. as if heard from a telephone receiver, to totally distorted, as if produced by an early 90’s video game console.

3. **Pitchshifter:** A frequency domain pitch shifter takes a Fast-Fourier transformed signal as an input and shifts the peaks found in the signal along the frequency access to transpose the audio (“gizmo~ Reference,” 2014). In practice, a gradual pitch shift is almost unbearable when applied to most music, so we implemented a discrete pitch shift transposing the audio up a major third, perfect fifth, and then octave when the user has remained on a time wasting website for three minutes, six minutes, and ten minutes respectively.

While the application has not been tested yet with users, we received IRB Approval (Protocol #15-0729) to run a multi-week field study in which users’ relevant internet browsing is monitored as they run the application with audio features disabled and enabled. A couple technical bugs must be ironed out, including one in which Mac security settings prevent the application from automatically resetting audio outputs. Sidestepping these limitations helped inform the design of the experiments described in Sections IV-VIII.

**Evaluation In Regards to Weiser and Brown’s Three Criteria Identifying Calm Technology**

1. **Movement:** Audio Poke takes advantage of the user’s music which, as described above, already moves between periphery and center of focus. Through altering the music, at first imperceptibly and later disruptively, we seek to nudge the music out of the periphery once it becomes pertinent, but most of the time we hope that the user detects the change on his/her own
naturally. The study described later provides an initial prediction of the effectiveness of this approach. Furthermore, we implemented Audio Poke for the Mac toolbar meaning it is largely hidden after an initial setup phase.

2. **Reach:** Audio Poke extends *peripheral reach* through taking an important consideration, whether the user is on a website he/she should be on, and moving it to an audio stream.

3. **Temporality:** Audio Poke relays a sense of familiarity through using user-selected music as its mode of communication. Similarly to the Sonification Station, it helps a user position himself/herself in time by altering audio gradually rather than communicating in discrete events.

**IV. EXPERIMENTAL DESIGN**

We designed a two-part experiment to study whether ambient audio alerts (involving altering background music or sounds) are feasible and may even be more effective than traditional audio alerts (e.g. ringtones) at providing workers with knowledge about external information while reducing the distraction from their primary task. As our motivation is to understand whether blended sonification of music may be a substitute for traditional notifications, the first experiment places blended sonification as a background notification in a work environment while the second experiment asks listeners to identify and describe their experiences while listening to blended sonifications. We have identified the following questions:

**Research Question 1:** How noticeable or disruptive are audio alerts in their role of providing information to a user? This is addressed in both experiments.

**Research Question 2:** Do ambient audio alerts provide any benefits to worker productivity? For example, do they allow the user to take natural stopping points when they switch tasks while at work? This is addressed in Experiment 1.

**Research Question 3:** How do people perceive and respond to different kinds of ambient audio alerts? Specifically, what audio effects are noticeable or annoying and how does the original, unaltered audio sample, the base for the blended sonification, affect perception? This is addressed slightly in Experiment 1 and primarily in Experiment 2.
Experiment 1: Task Productivity

Experiment 1 is meant to simulate a work environment in which workers have a primary job to do, but may also be receiving and responding to notifications, work-related or not. Currently, workers may receive notifications in the form of a ringtone or phone vibration, or they may elect to disable notifications entirely and check for external information at stopping points. Experiment 1 includes these two traditional methods of gathering external information, and introduces participants to a blended sonification of music.

A. Design

The experiment is a 1x3 within-subjects design in which the variable is a type of alert that participants hear. Participants work simultaneously on a primary task that stretches to fill the experiment and a secondary task that requires periodic monitoring. The secondary task may or may not be signalled by an audio alert depending upon the alert condition. Through the experiment, participants listen to music, sections of Music for 18 Musicians by Steve Reich. The three audio alert conditions are as follows:

1. **No Alert**: Participants are not notified when the secondary task becomes available. In other words, they must periodically stop working on the primary task and check if the secondary task has become available.

2. **Standard Alert**: A brief audio alert, similar to a standard text message or email tone, is sounded once the secondary task becomes available.

3. **Ambient Alert**: Over the course of one minute, the background music is altered with an effect that increases over the course of 30 seconds and then decreases over the course of 30 seconds. The effect selected for the experiment is a low-pass filter, identical to the effect used in Audio Poke described in Section III. As the level of effect is strengthened, the cutoff value of the filter decreases reducing the high frequencies damping and softening the audio quality.

B. Hypothesis
Given a primary task (e.g. transcribing a text file) and a secondary task (e.g. moving to press a button) indicated by an audio alert, we hypothesized that an ambient audio alert (e.g. blended music sonification) would enable workers to pause at a more natural stopping point in their primary task yielding an increase in focus and productivity. We measured the effect with a count of total words typed in a transcription task, a count of where breakpoints occurred in the transcriptions, and subjective responses to the alert conditions.

C. Implementation

We developed an application to automate this experiment utilizing four tools.

1. **Google Doc**: As a primary task, participants transcribed a series of images into a Google Doc shared between the participant and the researcher (Figure 4). The images were screenshots of David Sedaris essays published in the New Yorker magazine preventing text from being copied/pasted. We chose David Sedaris essays because they are entertaining, motivating to progress on, and all composed in a similar style. We selected a transcription for the primary task as it includes numerous stopping points, can expand to fill the available time in the experiment, and does not utilize any audio. Google Drive was selected as the transcribing platform due to the functionality of the Drive API (“Google Drive REST API,” 2016) which enabled us to easily download the participant’s progress at any time. The ordering was randomized for each experiment.
2. **Mira**: The secondary task was meant to mimic many real world tasks that require periodic checking such as responding to a time-sensitive email, checking if a bus has arrived, or pulling a batch of cookies out of the oven. It was conceived as a simple series of buttons to be pressed on an iPad (Figure 5) and implemented using the software Mira (Tarakajian, 2016). Mira enables a Max patch to be viewed and accessed on an iPad via wifi or USB. Triggering the patch on either the computer it is stored on or the iPad will alter the state on both devices in real time.
The secondary task required two button presses to complete and provided audio and visual feedback to the participant upon success. We made the decision to use two buttons in regards to the goals of the experiment: It was not crucial to know that the participant completed the secondary task. What was important however was when the participant paused progress on the primary task to attempt the secondary task. We added the top button to track every occasion the participant stopped working on the primary task.
3. **Max MSP**: The researcher ran the experiment via an application built in Max MSP (Figure 6). The application generated four alert/secondary task cycles with random intervals to fill each twelve-minute trial, selected a random musical segment and alert type before each trial, streamed the audio to the participant’s headphones via bluetooth, applied the low-pass filter effect during the ambient condition, and communicated with the other components powering the experiment. The interface enabled researchers to remotely monitor the current progress of each trial.

![Figure 6: Max MSP interface used by researchers](image)

4. **Python**: A script, running in the background through the course of the experiment, intercepted messages from Max MSP via the OSC protocol (Wright, 2002). It logged experiment events for later analysis and utilized the Drive API to download PDFs of the participant’s progress through the experiment.

**D. Data Analysis**

The experiment was evaluated in two ways: data collected while the experiment ran, and a subjective questionnaire filled out upon completion of the experiment (Appendix). Automatic data collection included a log file of all experimental events (Figure 7) with timestamps for the experiment’s start and finish, the availability of the secondary tasks, and the instances in which the user initiated the secondary task. In addition, PDFs of the transcription were downloaded automatically each time a participant initiated the secondary task. After the experiment concluded, researchers grouped the PDFs into three categories based on the participant’s most recent progress.
Either a word was left incomplete, a sentence was left incomplete, or a sentence was completed. Additionally, researchers counted the total number of words in transcriptions downloaded automatically at the end of trial. In the count, we regarded words as separated by spaces, hyphens, or other punctuation and did not take spelling into account. Using automatic data collection, we aimed to compare three measures across participants:

1. **Overall Progress**: Did the type of alert correspond with overall progress on the transcription task? (Research Question 2)

2. **Responsiveness to Alert**: After the alert was sounded, how long did participants take to initiate the secondary task? Did the participant ever fail to complete the secondary task during its window of availability? (Research Questions 1, 2)

3. **Work Interruption**: How often did participants pause and check whether the secondary task was available? When the participant moved to initiate the secondary task, where has he/she stopped in the middle of the transcription? In the middle of a sentence? In the middle of a word? Stopping at the end of a sentence represents a more natural break in workflow that we hoped ambient notifications would encourage. (Research Question 2)

The questionnaire included Likert scale measures for each condition including distraction, noticeability, annoyingness, and perceived level of success, as well as open-ended opportunities to describe overall responses to the alerts (Appendix).

**Experiment 2: Perception of Audio Effects**
Experiment 2 is similar to a hearing test. Before the experiment, researchers created a collection of minute-long audio samples from music of various genres and modified some samples with contrasting audio effects. During the experiment, participants listened to each sample and indicated when they heard an effect. They were asked to identify what they heard and whether they found it noticeable or annoying. The goals of this experiment are tied to Research Question 3.

A. Design

The experiment was a 1x5 within subjects design in which the variable was an effect applied to music by researchers. Participants heard an identical, randomly ordered set of ten audio files. Each audio file had one of four effects applied to it or it did not have an effect applied. All audio effects gradually increased for a duration and then decreased for an equivalent duration. Participants indicated during each audio sample where they perceived an effect beginning or ending.

In this work, we are motivated to design effective ambient notifications that accommodate most users and as such, we aim to understand the range in individual differences of threshold detection and subjective response. The audio samples selected reflect an eclectic mix of genres and instrumentation, and some are more well-known than others. Effects were also intended to be wide-varying with some directly altering the timbre of the original audio, and others manipulating the audio more subtly.

B. Hypothesis

Since this research is nascent and designed to begin to understand the characteristics and potential tradeoffs between effects, we did not compose a formal hypothesis. We primarily aimed to understand the differences between effects. That said, we predicted a positive correlation between perceived annoyingness and noticeability as well as a positive correlation between musical experience and accuracy across all audio effects.

C. Implementation

An application built using Max MSP, as with Experiment 1 and the two prototypes, randomly selected an audio sample to play to the participant and included a button that may be
pressed when the participant perceives the existence of an effect (Figure 8). Buttons were made inactive or active to the participant depending on his/her progress in the experiment. As a result, Experiment 2 was fully automated without researcher action.

![Experiment 2 interface](image)

Figure 8: Experiment 2 interface used by participant while audio sample is playing

We selected four effects to apply to some of the audio samples. To create the samples and apply the effects we used the software Logic Pro X (Figure 9). First, researchers added an original audio file into the software interface, and isolated a minute-long chunk of the audio removing the rest of the original file. Then, using prebuilt effects in Logic, researchers applied an effect to a subsection of the sample. The effect increased in magnitude for half of the subsection and then decreased in magnitude for half of the subsection, both on a linear curve. The start of the subsection and the length of the subsection differed for each audio segment. Audio was rendered ahead of the experiment to ensure that participants heard identical audio samples and because realtime audio manipulation was unnecessary for the experimental design. The audio effects we chose are as following:

1. Signal Distortion: The quality of the audio becomes reduced and may sound buzzy or distorted.
2. Low Pass Filter: High frequencies are removed causing the audio to sound lower and muffled.
3. Echo Effect: The audio quickly echos once.
4. Panned Tremolo: The audio moves back and forth between the two headphones.

![Audio samples rendered in Logic Pro X. Green lines indicate an effect fading in and out](image)

Ten audio samples were selected representing a range of styles and instrumentation (Table 1). From feedback provided during pre-tests, we judged four audio effects to be particularly difficult to perceive. The audio sample *Cello Song* uses layered instrumentation that may be mistaken for an effect, the echo effect in *Invention No. 3* is set to the same tempo as the music and may be perceived as natural concert hall reverb, the unpredictable melody and natural panning of the piano in *A Lark* may mask the panned tremolo, and *Hannibal* is composed using a panning effect and heavy reverb effect that may be mistaken for effects applied by researchers.

**Table 1: Audio Samples Used in Experiment 2**

<table>
<thead>
<tr>
<th>Audio Sample</th>
<th>Style</th>
<th>Effect Applied</th>
<th>Difficulty</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Freddie Freeloader</em> by Miles Davis</td>
<td>Jazz Sextet</td>
<td>Signal Distortion</td>
<td>Easy</td>
</tr>
<tr>
<td><em>Cello Song</em> by Nick Drake</td>
<td>Folk</td>
<td>Low Pass Filter</td>
<td>Hard</td>
</tr>
<tr>
<td><em>Thinkin Bout You</em> by Frank Ocean</td>
<td>R&amp;B</td>
<td>Echo Effect</td>
<td>Easy</td>
</tr>
<tr>
<td><em>Fire’s Highway</em> by Japandroids</td>
<td>Garage Rock</td>
<td>Low Pass Filter</td>
<td>Easy</td>
</tr>
<tr>
<td><em>Invention No. 3</em> by Bach</td>
<td>Classical, Solo Piano</td>
<td>Echo Effect</td>
<td>Hard</td>
</tr>
<tr>
<td><em>Lenny</em> by Stevie Ray Vaughn</td>
<td>Blues, Electric Guitar</td>
<td>Panned Tremolo</td>
<td>Easy</td>
</tr>
<tr>
<td><em>Hannibal</em> by Caribou</td>
<td>Electronic</td>
<td>None</td>
<td>Hard</td>
</tr>
<tr>
<td><em>A Lark</em> by Fred Hersch</td>
<td>Jazz, Solo Piano</td>
<td>Panned Tremolo</td>
<td>Hard</td>
</tr>
</tbody>
</table>
Yesterday by the Beatles | Pop | Signal Distortion | Easy
Violin Sonata No. 2 by Bach | Classical, Solo Violin | None | Easy

D. Data Analysis

Each time a participant detected that an effect had begun or ended, the experiment application automatically logged an exact timestamp in milliseconds. With these timestamps, we counted instances of false positives, e.g. when the participant hears an effect though one has not been applied, and instances of false negatives, e.g. when an effect has been applied though the participant has not perceived it. We also compared the timing data between when the effects actually occurred and when the participants detected them.

The following shows results in the format collected during the experiment (Table 2). In both tables, a time of 0 indicates the start of an audio sample. In the left table, the two times following a 0 indicate the time an effect is applied and then the time the effect is removed. In the right table, the two times represent when the participant perceives the start and end of an effect. In some cases, the participant did not notice the end of an effect, shown by only a single number following a 0.
Table 2: Automatic Data Collection for Experiment 2

<table>
<thead>
<tr>
<th>ID</th>
<th>Audio Sample</th>
<th>Effect Time (ms)</th>
<th>ID</th>
<th>Audio Sample</th>
<th>Event Time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Act. 1</td>
<td>1</td>
<td>0</td>
<td>Khr. 1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Act. 1</td>
<td>1</td>
<td>14000</td>
<td>Khr. 1</td>
<td>1</td>
<td>24826</td>
</tr>
<tr>
<td>Act. 1</td>
<td>1</td>
<td>38000</td>
<td>Khr. 1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Act. 2</td>
<td>2</td>
<td>0</td>
<td>Khr. 2</td>
<td>3</td>
<td>34141</td>
</tr>
<tr>
<td>Act. 2</td>
<td>2</td>
<td>28000</td>
<td>Khr. 2</td>
<td>2</td>
<td>41879</td>
</tr>
<tr>
<td>Act. 2</td>
<td>2</td>
<td>44000</td>
<td>Khr. 2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Act. 3</td>
<td>3</td>
<td>0</td>
<td>Khr. 3</td>
<td>3</td>
<td>27036</td>
</tr>
<tr>
<td>Act. 3</td>
<td>3</td>
<td>10000</td>
<td>Khr. 3</td>
<td>3</td>
<td>34152</td>
</tr>
<tr>
<td>Act. 4</td>
<td>4</td>
<td>0</td>
<td>Khr. 4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Act. 4</td>
<td>4</td>
<td>6000</td>
<td>Khr. 4</td>
<td>4</td>
<td>18194</td>
</tr>
<tr>
<td>Act. 4</td>
<td>4</td>
<td>50000</td>
<td>Khr. 4</td>
<td>4</td>
<td>45446</td>
</tr>
<tr>
<td>Act. 5</td>
<td>5</td>
<td>0</td>
<td>Khr. 5</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Act. 5</td>
<td>5</td>
<td>6000</td>
<td>Khr. 6</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Act. 5</td>
<td>5</td>
<td>40000</td>
<td>Khr. 6</td>
<td>6</td>
<td>42446</td>
</tr>
<tr>
<td>Act. 6</td>
<td>6</td>
<td>0</td>
<td>Khr. 6</td>
<td>6</td>
<td>51864</td>
</tr>
<tr>
<td>Act. 6</td>
<td>6</td>
<td>30000</td>
<td>Khr. 6</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Act. 6</td>
<td>6</td>
<td>50000</td>
<td>Khr. 7</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Act. 7</td>
<td>7</td>
<td>0</td>
<td>Khr. 8</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Act. 8</td>
<td>8</td>
<td>0</td>
<td>Khr. 8</td>
<td>8</td>
<td>29528</td>
</tr>
<tr>
<td>Act. 8</td>
<td>8</td>
<td>10000</td>
<td>Khr. 8</td>
<td>8</td>
<td>42188</td>
</tr>
<tr>
<td>Act. 8</td>
<td>8</td>
<td>50000</td>
<td>Khr. 9</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Act. 9</td>
<td>9</td>
<td>0</td>
<td>Khr. 9</td>
<td>9</td>
<td>12423</td>
</tr>
<tr>
<td>Act. 9</td>
<td>9</td>
<td>2000</td>
<td>Khr. 10</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Act. 9</td>
<td>9</td>
<td>38000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Act. 10</td>
<td>10</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Furthermore, after each audio sample, participants filled out a short survey pertaining to what they heard (Appendix). The survey included objective measures, e.g. whether the participant
correctly classified the effect they heard, and subjective measures including Likert scales with subjects such as Annoyingness and Noticeability.

V. METHODS

The Study Received IRB Approval on September 9, 2016 (Protocol #16-0588), and was run between October 13, 2016 and October 28, 2016. All experiments took place in the Superhuman Computing Lab with a researcher present and lasted, on average, one hour and fifteen minutes. A researcher read from a script to improve consistency across experiments and all participants were given a random pseudonym dissociating experimental data with true identities. Before actual trials began, both experiments were run with members of the Superhuman Computing Lab and slightly adjusted due to feedback. The experiment was advertised via emails sent out to students in the College of Engineering, the College of Music, a campus Hackerspace, and through word-of-mouth.

First, participants were asked to fill out the consent form and brief pre-survey (Appendix), and were provided an overview of the experiment and its motivations (Figure 10). Then, the researcher gave a detailed description of Experiment 1 including a demo of the iPad. As many elements as possible were randomized during the experiment including the ordering of documents to transcribe, segments of music to listen to, and alert conditions. Experiment 1 took, on average, 50 minutes to run including three 12-minute trials, descriptions and demos of the relevant alert type before each trial, and time to complete the post-survey. Participants wore bluetooth headphones and completed the experiment on a desktop computer placed in the lab, while a researcher faced in the other direction and monitored the experiment on a laptop. After completion of Experiment 1, participants were provided an overview of Experiment 2 and given a demo of an audio sample and each of the four possible effects applied to that sample. Experiment 2 lasted 15 minutes on average and was completed on a laptop with a “cheatsheet” nearby detailing the four possible audio effects. Finally, upon completion of both experiments, participants wrote their email addresses on paper to receive $10 compensation for their time.
Participants

Nine participants completed both experiments while one additional participant completed only Experiment 2. Participants varied in gender (Male = 5, Female = 4, Other = 1) and musical experience (None or Little = 3, Some = 4, Significant = 3). They varied less so in occupation (Student = 9, Musician = 1) and age (Between 18 and 24 = 7, Between 25 and 34 = 3). On a scale of 1-5, participants enable cell phone audio notifications a small part of the time ($M = 2.4, SD = 1.17$), while they keep vibration enabled most of the time ($M = 3.8, SD = 1.4$). On the computer, they disable audio notifications most of the time ($M = 2.0, SD = 1.33$) and listen to music part of the time ($M = 3.1, SD = 1.29$). Each participant uses a wide range of mobile services that deliver notifications including texts, Snapchat, calendar, Instagram, and email, and a smaller range of desktop applications including Facebook, Slack, and Gmail. They listen to an eclectic range of music as well including Progressive Rock, Hip Hop, Classical (Music without Words), and Indie Rock.

VI. RESULTS
This section presents our analysis of data collected in both experiments beginning with objective measures and transitioning to Likert Scales and subjective responses collected from surveys.

**Experiment 1**

Dependent upon the alert condition, we considered the number of times the participant checked if the secondary task was available (Figure 11), the number of times the participant completed the secondary task, and the total number of words the participant typed. Mauchly's test showed a violation of sphericity against Alert Type ($W(2) = 0.38, p < 0.05$). We ran one-way repeated-measure ANOVA with Greenhouse-Geisser correction ($\varepsilon = 0.62$) revealing a significant effect of Alert Type on the number of times participants checked if the secondary task was available ($F(1.23, 9.85) = 23.77, p < 0.001, \text{partial } n^2 = 0.75$). In the following bar graph, a lower number of checks indicates a lower number of times that the participant stopped working on the primary task. Since the secondary task occurred four times during the experiment, the ideal number of checks is four.

![Alert Type Vs. Number of Checks](image)

Figure 11: Bar graph showing type of alert against number of times the participant engaged the secondary task
We found no significant effect of Alert Type on either the number of secondary tasks successfully completed (Gradual $M = 3.78$, $SD = 0.67$; None $M = 3.56$, $SD = 0.72$; Standard $M = 3.89$, $SD = 0.33$) or on the total number of words typed by the participants (Gradual $M = 551.7$, $SD = 135.2$; None $M = 568.9$, $SD = 140.9$; Standard $M = 564.8$, $SD = 140.1$).

We tracked for each alert, the time that the participant took to complete the secondary task (Figure 12). This is essentially the participant’s time away from the keyboard and was measured automatically as the difference in seconds between the secondary task becoming available and the participant pressing the blue button to complete the secondary task. The participants had up to 60 seconds to complete the secondary task. A short time indicates that participants felt obligated to complete the secondary task directly after it becoming available, while a large standard deviation indicates wide differences between participants completing the secondary task. 7 instances out of 108 total alerts occurred in which a participant failed to complete the secondary task (Gradual = 2, None = 4, Standard = 1). These were considered “misses” and were removed from the data. Mauchly's test showed a violation of sphericity against Alert Type ($W(2) = 0.24$, $p < 0.05$). We ran one-way repeated-measure ANOVA with Greenhouse-Geisser correction ($\varepsilon = 0.57$). It revealed a significant effect of Alert Type on the length of time in which participants took to complete the secondary task ($F(1.14, 9.01) = 32.16$, $p < 0.001$, partial $\eta^2 = 0.80$).
Additionally, we tracked the stopping points of the participant each time the secondary task was initiated (Table 3). This was automatically measured when the participant pressed the top button of the secondary task. Since the secondary task was initiated significantly more during the No Alert condition, overall counts are much higher than the other conditions. Researchers manually labelled every downloaded PDF observing the participant’s most recent progress as either having left the current word incomplete (e.g. “th” for “the”), having left the current sentence incomplete (e.g. “I went” for “I went to the store.”), or having completed the sentence. PDFs included data on the time they were downloaded, the experiment they were downloaded during, and the participant who worked on them. We ran Cramer’s V to determine if there is an association between Alert Type and Participant’s Current Progress and found a relatively weak association (Cramer’s V = 0.21).
Table 3: Crosstab Table Showing Participant’s Progress by Alert Type

<table>
<thead>
<tr>
<th>Alert Type</th>
<th>Level of Incomplete</th>
<th>Character Raw (%)</th>
<th>Word Raw (%)</th>
<th>Sentence Raw (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard (n = 36)</td>
<td></td>
<td>6 (16.7%)</td>
<td>17 (47.2%)</td>
<td>13 (36.1%)</td>
</tr>
<tr>
<td>Gradual (n = 39)</td>
<td></td>
<td>6 (15.4%)</td>
<td>27 (69.2%)</td>
<td>6 (15.4%)</td>
</tr>
<tr>
<td>None (n = 81)</td>
<td></td>
<td>3 (3.7%)</td>
<td>43 (53.1%)</td>
<td>35 (43.2%)</td>
</tr>
</tbody>
</table>

Surveys collected after Experiment 1 revealed participants’ reactions to each condition. We measured distraction level of the alert, level of noticeability, level of annoyingness, usefulness of the alert, how successful the participant felt he/she was, and how hard the participant felt he/she had to work (Figure 13, Table 4).
Figure 13.1: Likert scale responses following Experiment 1 for gradual alert condition
Figure 13.2: Likert scale responses following Experiment 1 for no alert condition
Figure 13.3: Likert scale responses following Experiment 1 for standard alert condition

Table 4: Likert Scale Responses for Each Alert

<table>
<thead>
<tr>
<th>Measure</th>
<th>Gradual Alert Mean (SD)</th>
<th>No Alert Mean (SD)</th>
<th>Standard Alert Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distraction</td>
<td>2.8 (1.5)</td>
<td>2.8 (2.1)</td>
<td>2.7 (1.2)</td>
</tr>
<tr>
<td>Noticeability</td>
<td>3.2 (1.2)</td>
<td>1.0 (0.0)</td>
<td>4.7 (0.5)</td>
</tr>
<tr>
<td>Annoyingness</td>
<td>1.6 (1.3)</td>
<td>3.2 (1.8)</td>
<td>2.7 (1.1)</td>
</tr>
<tr>
<td>Usefulness</td>
<td>3.2 (1.5)</td>
<td>1.0 (0.0)</td>
<td>4.8 (0.5)</td>
</tr>
<tr>
<td>Success</td>
<td>3.9 (0.9)</td>
<td>3.0 (1.0)</td>
<td>4.2 (0.8)</td>
</tr>
<tr>
<td>Difficulty</td>
<td>3.8 (1.0)</td>
<td>4.1 (0.9)</td>
<td>2.8 (1.1)</td>
</tr>
</tbody>
</table>
Four participants expressed confusion about the distraction and annoyingness questions in the No Alert condition while filling out the post-survey. The wide range of responses stems from two ways of answering: It is not at all distracting/annoying because no audio notification interrupts primary workflow, but having to constantly consider the Secondary Task is distracting/annoying.

Overall in the post-survey, five out of nine participants rated feeling most productive during the gradual alert condition, while three identified the regular alert, and one identified no alert. Seven out of nine participants rated the no-alert condition as most greatly inhibiting productivity, while two identified the gradual alert. Seven participants identified the regular alert as their “favorite” while two selected the gradual alert.

**Experiment 2**

As participants listened to each audio sample, they could have made one of two types of errors, first a Type 1 (false positive) error if they indicated during audio playback that they heard an effect when researchers had in fact not altered the audio sample. Or, they could make a Type 2 (false negative) error if they did not indicate hearing an audio effect, but researchers in fact altered the audio. There were two opportunities for participants to make a Type 1 error, and eight opportunities for participants for make a Type 2 error. Overall, these errors were rare with most participants correctly detecting the presence of an effect or lack thereof in each audio sample (Figure 14).
Figure 14: Errors participants made in Experiment 2, grouped by musical training

After the audio sample completed, participants were asked to select whether they heard Signal Distortion, Low Pass Filter, Echo, Panned Tremolo, or No Effect. We counted the number of correctly identified effects per participant and once again grouped the counts based on musical experience (Figure 15). In some cases, even if the participant made a Type 1 error, he/she realized at the end of audio playback that no effect had been applied and still made a correct identification. In only one case did this occur for a Type 2 error: The participant correctly indicated hearing an echo effect during audio playback, but afterwards decided it was natural reverb and indicated “No Effect” in the survey.
During audio playback, we calculated the time delay in milliseconds between the addition or removal of an audio effect and the participant’s detection (Figure 16). A positive value indicates that the participant detected an effect change some time after it was applied/removed, while a negative value indicates that the participant detected an effect change before it was applied/removed. Values that are closer to 0 indicate that participants detected an effect close to when it was added or removed. Values with smaller deviations indicate that participants detected an effect around the same time. Occurrences in which the participant did not hear an effect, and thus did not provide any timing data, were removed from the dataset.
Figure 16.1: Participant detection times for audio samples when an audio effect was added
After each audio segment, participants were asked how noticeable and annoying they felt the effect was. Participants were only asked to rate these factors if an effect was actually detected (Table 5). As a result, the sample size varies across audio segments.

Table 5: Likert Scale Responses and Identification Data Taken from Experiment 2 Survey

<table>
<thead>
<tr>
<th>Song</th>
<th>Effect</th>
<th>N</th>
<th>Noticeability Mean (SD)</th>
<th>Annoyingness Mean (SD)</th>
<th>Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Distortion</td>
<td>10</td>
<td>4.6 (0.70)</td>
<td>4.1 (1.10)</td>
<td>10/10</td>
</tr>
<tr>
<td>2</td>
<td>Low Pass</td>
<td>10</td>
<td>4.8 (0.42)</td>
<td>1.8 (0.63)</td>
<td>10/10</td>
</tr>
<tr>
<td></td>
<td>Effect</td>
<td></td>
<td>Likert Scale Ratings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>------------</td>
<td>---</td>
<td>----------------------</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>3</td>
<td>Echo</td>
<td>8</td>
<td>3.4 (1.06)</td>
<td>2.1 (0.83)</td>
<td>8/10</td>
</tr>
<tr>
<td>4</td>
<td>Low Pass</td>
<td>10</td>
<td>4.6 (0.70)</td>
<td>2.1 (0.88)</td>
<td>10/10</td>
</tr>
<tr>
<td>5</td>
<td>Echo</td>
<td>6</td>
<td>2.5 (0.84)</td>
<td>2.5 (1.22)</td>
<td>5/10</td>
</tr>
<tr>
<td>6</td>
<td>Tremolo</td>
<td>10</td>
<td>4.3 (0.68)</td>
<td>3.2 (1.75)</td>
<td>10/10</td>
</tr>
<tr>
<td>7</td>
<td>None</td>
<td>4</td>
<td>2.25 (0.96)</td>
<td>3.0 (1.15)</td>
<td>6/10</td>
</tr>
<tr>
<td>8</td>
<td>Tremolo</td>
<td>9</td>
<td>3.7 (0.71)</td>
<td>2.9 (1.27)</td>
<td>8/10</td>
</tr>
<tr>
<td>9</td>
<td>Distortion</td>
<td>10</td>
<td>4.9 (0.31)</td>
<td>4.3 (0.82)</td>
<td>10/10</td>
</tr>
<tr>
<td>10</td>
<td>None</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>9/10</td>
</tr>
</tbody>
</table>

The following chart plots the averaged Likert Scale Ratings of Annoyingness against the averaged Likert Scale Ratings of Noticeability (Figure 17). Each point reflects the responses of participants for an audio sample. Since each audio effect was applied to two audio samples, there are two points for each audio effect. The top right corner of the plot indicates an effect was simultaneously noticeable and annoying, while the bottom right corner is the sweet spot: The effect was found to be noticeable, but not annoying.
VII. DISCUSSION

Experiment 1

Unfortunately, the results from Experiment 1 are not strong enough to support our hypothesis. The Gradual Alert neither enhanced productivity, which would have shown by increased primary task output and secondary task completion, or functioned in the way we envisioned it would as a calm technology. Participants did not benefit from the gradual change in background music and instead likely treated the Gradual Effect similarly to the Standard Effect. One participant wrote about the Gradual Alert that he was, “constantly worried [he] would miss it, so [he] was always thinking about it.” The standard deviations in secondary task completion of both conditions are also similar (Standard: \( M = 6.5, SD = 5.4 \), Gradual: \( M = 14.8, SD = 7.2 \)). A successful calm notification should produce a larger standard deviation resulting from individual and moment-to-moment differences. Participants almost never waited past 30 seconds to complete.

Figure 17: Likert scale ratings for Noticeability and Annoyingness, grouped by audio effect
the secondary task even though the Gradual Alert was meant to provide temporal comfort for the full minute the secondary task was available. Additionally, the Gradual Alert showed no difference in supporting participants to find natural breaks as they shifted concentration away from the primary task. It produced a similar distribution in stopping places as the Standard Alert.

Even, though we could not verify our hypothesis, we replicated results from past studies, including a general frustration when all notifications are removed from the workplace (Iqbal & Horvitz, 2010). We also uncovered interesting results that may necessitate further study of ambient notifications. While one person wrote that the Gradual Alert would be useful for nothing and “hated it,” most participants brainstormed possible uses. For example, one wrote that it would be useful “when the second task is non-critical, such as a personal conversation.” Another explained it could be helpful for “some type of interruption that is of the lesser urgent variety. It may be a good signal for stopping time or long-term task switching.” It is also no small feat that participants were introduced to blended sonification with a 20-second sample, and performed similarly to the Standard Effect, a condition matching the types of notifications many have been exposed to for years. Two participants even indicated as such, stating they may like the gradual alert better if they had used it for a more sustained time period.

If we were to conduct this experiment again, there are other measures that may help answer our research questions. First, we should measure the time it takes participants to get back on the main task after initiating the secondary task. An ideal calm technology should facilitate this transition reducing the time it takes to multi-task. Second, we should measure the typing rate of participants across the experiment. If any single condition reduces participants’ typing rate, it could indicate difficulty in using the alert causing them to divide their attention. Third, we should run experiments with participants’ own music. Detectability was a concern for the gradual alert condition, and one factor may be that participants listened to unfamiliar music. They had a harder time noticing when the music was altered without a fundamental knowledge of what the music sounds like unaltered. They may find the gradual alert more detectable with music they are already familiar with.

Future studies may have more success should they adapt the experimental design to best promote the hypothesized benefits of ambient audio. For example, participants should work on a
more cognitively demanding primary task in which the consequences for switching tasks without care are more severe. For example, instead of a transcription, the task could be a Sudoku puzzle or previously shown to be especially affected by notifications such as finding items from a long list (Cutrell et al., 2001). A longer field study may also yield results where participants have the chance to learn and live with ambient notifications. Finally, participants could be asked to bring in the music they already enjoy and put on in the background while they work.

Outside of our experimental questions, we uncovered wide differences among the participants with the data collected. The total number of words typed in a twelve-minute session ranged from 374 to 804. Some participants demonstrated learned competency in working under the distraction of notifications. For example, one participant who showed longer times ($M = 8.8$ seconds) than average ($M = 6.5$ seconds) to complete the secondary task during the Standard Alert condition, also never left a word incomplete to initiate the secondary task. This same participant included a longer list than others of mobile and desktop applications that send notifications. On the other end of the spectrum, another participant showed shorter times ($M = 2.5$ seconds) to complete the secondary task during the Standard Alert condition, and left a word incomplete on three occasions during the entire experiment. Future research on ambient audio notifications should take these dramatic differences into account perhaps adapting and increasing the time it takes to fade the audio in as the participant learns to use it.

**Experiment 2**

Results from Experiment 2 provide us with a good foundation for exploring the application of audio effects in future work. Promisingly, all participants demonstrated an ability to learn and accurately detect and recognize a series of contrasting audio effects. Participants with a high degree of musical training proved slightly more adept at recognizing the presence/absence of an audio effect. They made less Type I and Type II errors ($M = 0.3$) than participants with some musical training ($M = 1.5$) and participants with little or no musical training ($M = 2.0$). They also correctly identified more audio samples correctly out of ten ($M = 9.3$) than participants with some musical training ($M = 8.5$) and participants with little or no musical training ($M = 8.3$). A larger sample size of various musical backgrounds may strengthen this result. Three effects, distortion (20/20
identifications), low pass filter (20/20 identifications), and tremolo (18/20 identifications) proved easier to detect than echo (13/20 identifications). It is worth noting that participants used a low pass filter effect for experiment 1, and may have already been trained to detect it.

We predicted that noticeability would positively correlate with annoyingness. This proved true in the case of distortion which rated high in both measures and even prompted some emotional responses: One participant indicated he was distracted since he was “laughing at how Freddie Freeloader sounded with distortion.” This was not true for the tremolo effect which rated high for noticeability and medium for annoyingness, or the low pass filter which showed the best of both worlds maximizing noticeability and minimizing annoyingness. The low pass filter has its own limitations though: Occurrences of it resulted in the largest standard deviations of effect ending times ($SD2 = 3185 \text{ ms}, SD4 = 3690 \text{ ms}$) and it is likely less accessible for users with some hearing loss who have difficulty detecting higher frequencies. The distortion effect may be advantageous in especially critical notifications since it provokes strong emotions and resulted in a relatively small standard deviation in both audio samples for initial detection time ($SD1 =1879 \text{ ms }, SD9 = 2144 \text{ ms}$), and ending detection time ($SD1 = 1230, SD9 = 1884$). Finally, even though it was harder to detect, an echo effect may still be worth exploring. The participants who detected it found it not to be annoying, and one participant even specified it as her favorite. Other features, such as feedback or level of the original signal, could be manipulated as the echo’s amplitude grows enhancing its capability to be discerned.

Two audio samples yielded especially interesting results. First in Sample 5 taken from Bach’s Third Invention, participants had extra difficulty detecting an echo effect. The audio sample is a piano solo in which the echo effect results in piano keys layering on top of and harmonizing with the original performance. Only five participants detected the effect and must have gone back and forth between whether they heard natural or artificial audio as indicated by the large range of detection times ($M = 19808 \text{ ms}, SD = 9691 \text{ ms}, Range = 23758 \text{ ms}$). Even though this is an entirely acoustic recording, the artificial effect blended in enough to cause confusion.

A second interesting case to dive into is Sample 7 taken from Hannibal by the artist Caribou. This electronic track featured no additional audio effect applied by researchers, but was composed with numerous effects including heavy reverb similar to the echo effect, panning similar to the
panned tremolo effect, and some crunchy distortion. Participants were confounded by this sample: Six out of ten participants heard an effect during playback, though two indicated they made a mistake and identified the sample in the survey as having no effect. Another three however concluded the sample had a tremolo effect while one selected signal distortion. No advanced musicians made this Type 1 error. Sample 7 presents an inherent difficulty in using blended sonification over electronic music because electronic music already uses many audio effects, sometimes in clever, unexpected ways. In contrast, no participants confused Track 2, a garage rock sample, as having distortion applied likely because distorted electric guitars are such a commonplace facet of rock music. Future research should explore whether certain audio effects are more highly detectable over electronic music, whether people are better able to detect audio effects in familiar electronic music vs. unfamiliar electronic music, and whether people with musical backgrounds have an advantage. It may be predicted that advanced listeners (electronic artists or composers) are better suited to detect an applied effect since they may be able to think from the artist’s point-of-view and intuit when an effect fits within the artist’s compositional objective or when it is extraneous.

Another finding is that participants were much more accurate in detecting the ending of an effect (Means: -2020.2, 528.6, 322.1, -4075.3, 2813, 1858.7, -7202.7, -4092.4) than the beginning (Means: 10545.5, 4672.4, 8689.4, 8769.7, 19807.4, 9924.5, 17258.67, 10226.5). One reason for this is availability of detection times both before and after the effect has ended once the effect has reached imperceptibly low levels. This availability does not exist at the start of an effect where the participant would have to predict the future. Another possible cause is that participants had already heard and identified the effect by the point it is reaching its end. The task of identifying an effect has disappeared is less demanding than identifying an effect has begun. At the beginning the participant has no idea what effect there might be or whether one may occur at all. Future work in designing notification systems should explore how best to utilize this feature of ambient audio.

VIII. CONCLUSION
We presented novel contributions on the usage of ambient audio and introduced the application of blended sonification on music. Two prototype systems, Sonification Station and Audio Poke, show how ambient audio may integrate into our daily lives. Experiment 2 begins to build an understanding of how people perceive and respond to audio effects. Even though Experiment 1 did not produce strong evidence showing the benefit of ambient audio notifications, we believe that this preliminary research shows promise, and that future experiments may uncover possible advantages of ambient notification systems. Current methods for delivering external data are not compatible with Weiser and Brown’s criteria identifying a calm technology. Audio notifications often wrench our attention away from our most critical task. Yet, Removing notifications entirely increases the information we must consider rather than letting it remain in the periphery. If the technologies surrounding us are to evolve and become calm, new notification systems must continue to be introduced and explored.
References


Psychiatrists, 7(3), 231–239.

https://docs.cycling74.com/max7/maxobject/degrade~


Predicting Human Interruptibility with Sensors. ACM Transactions on Computer-Human
Interaction: A Publication of the Association for Computing Machinery, 12(1), 119–146.

https://docs.cycling74.com/max7/maxobject/gizmo~

https://developers.google.com/drive/v3/web/about-sdk

offices in the UK. Musicae Scientiae: The Journal of the European Society for the Cognitive

http://justgetflux.com/


Iqbal, S. T., & Bailey, B. P. (2008). Effects of Intelligent Notification Management on Users and


337–360.

https://selfcontrolapp.com


https://cycling74.com/products/mira


http://sonify.psych.gatech.edu/research/sonification_sandbox/


Appendix
Pre-Survey
SuperHuman Computing Lab - Audio Alert Study

* Required

1. **User ID** *
   (Provided by Researchers)

Questions

2. **List the mobile phone services or applications you use that automatically send you notifications.** *

   ........................................................................................................................................

   ........................................................................................................................................

   ........................................................................................................................................

   ........................................................................................................................................

   ........................................................................................................................................

3. **How often do you enable audio for phone notifications? (e.g. turn ringtone on)***

   *Mark only one oval.*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. **How often do you enable vibration for phone notifications?** *

   *Mark only one oval.*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5. List the computer services or applications you use that automatically send you notifications. *

6. How often do you enable audio for desktop notifications? *
   Mark only one oval.

   1  2  3  4  5
   Never ☐ ☐ ☐ ☐ ☐ Always

7. How often do you listen to music while you work? *
   Mark only one oval.

   1  2  3  4  5
   None of the time ☐ ☐ ☐ ☐ ☐ All of the time

8. What genres of music do you listen to when you work? *

9. What statement most closely describes your level of musical training? *
   Mark only one oval.

   ☐ No musical experience
   ☐ No formal training, but some experience playing an instrument
   ☐ Some formal training, e.g. private lessons, high school band/orchestra
   ☐ College degree in music / Professional musician
   ☐ Other: ......................................................................................................

Demographics
10. **Gender**

*Mark only one oval.*

- [ ] Female
- [ ] Male
- [ ] Prefer Not To Say
- [ ] Other: ............................................................................................................................

11. **Occupation**

............................................................................................................................

12. **Age**

*Mark only one oval.*

- [ ] 18 - 24
- [ ] 25 - 34
- [ ] 35 - 44
- [ ] 45 - 54
- [ ] 55 - 64
- [ ] 65 - 74
- [ ] 75 or greater

SUPERHUMAN COMPUTING LAB

at the University of Colorado Boulder
Post-Survey
SuperHuman Computing Lab - Audio Alert Study

* Required

1. **User ID** *
   (Provided by Researchers)

---

**Condition 1**

2. **What was the first type of Audio Alert you heard?** *
   
   *Mark only one oval.*
   - [ ] No Alert
   - [ ] Standard Alert
   - [ ] Gradual (Ambient) Alert

3. **How distracting was this alert?** *
   
   *Mark only one oval.*
   
<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all distracting</td>
<td></td>
<td></td>
<td></td>
<td>Very distracting</td>
</tr>
</tbody>
</table>

4. **How noticeable was this alert?** *

   *Mark only one oval.*
   
<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all noticeable</td>
<td></td>
<td></td>
<td></td>
<td>Very noticeable</td>
</tr>
</tbody>
</table>

5. **How annoying was this alert?** *
   
   *Mark only one oval.*

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all annoying</td>
<td>Very annoying</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6. How useful was this alert in telling you about the button? *  
*Mark only one oval.*  

1 2 3 4 5  
Not useful at all  □ □ □ □ □  Very useful

7. How successful were you in accomplishing what you were asked to do? *  
*Mark only one oval.*  

1 2 3 4 5  
Not at all successful □ □ □ □ □  Very successful

8. How hard did you have to work to accomplish your level of performance? *  
*Mark only one oval.*  

1 2 3 4 5  
Not hard at all □ □ □ □ □  Very hard

Condition 2

9. What was the second type of Audio Alert you heard? *  
*Mark only one oval.*  

□ No Alert  
□ Standard Alert  
□ Gradual (Ambient) Alert

10. How distracting was this alert? *  
*Mark only one oval.*  

1 2 3 4 5  
Not at all distracting □ □ □ □ □  Very distracting

11. How noticeable was this alert? *  
*Mark only one oval.*  

1 2 3 4 5  
Not at all noticeable □ □ □ □ □  Very noticeable
12. **How annoying was this alert?** *  
*Mark only one oval.*

1 2 3 4 5

Not at all annoying  |||  Very annoying

13. **How useful was this alert in telling you about the button?** *  
*Mark only one oval.*

1 2 3 4 5

Not useful at all  |||  Very useful

14. **How successful were you in accomplishing what you were asked to do?** *  
*Mark only one oval.*

1 2 3 4 5

Not at all successful  |||  Very successful

15. **How hard did you have to work to accomplish your level of performance?** *  
*Mark only one oval.*

1 2 3 4 5

Not hard at all  |||  Very hard

**Condition 3**

16. **What was the third type of Audio Alert you heard?** *  
*Mark only one oval.*

No Alert  |||  Standard Alert  |||  Gradual (Ambient) Alert

17. **How distracting was this alert?** *  
*Mark only one oval.*

1 2 3 4 5

Not at all distracting  |||  Very distracting
18. **How noticeable was this alert?** *

*Mark only one oval.*

1 2 3 4 5

Not at all noticeable  [ ] [ ] [ ] [ ] [ ] Very noticeable

19. **How annoying was this alert?** *

*Mark only one oval.*

1 2 3 4 5

Not at all annoying  [ ] [ ] [ ] [ ] [ ] Very annoying

20. **How useful was this alert in telling you about the button?** *

*Mark only one oval.*

1 2 3 4 5

Not useful at all  [ ] [ ] [ ] [ ] [ ] Very useful

21. **How successful were you in accomplishing what you were asked to do?** *

*Mark only one oval.*

1 2 3 4 5

Not at all successful  [ ] [ ] [ ] [ ] [ ] Very successful

22. **How hard did you have to work to accomplish your level of performance?** *

*Mark only one oval.*

1 2 3 4 5

Not hard at all  [ ] [ ] [ ] [ ] [ ] Very hard

**General Questions**

23. **During which alert do you feel you were the most productive?** *

*Mark only one oval.*

[ ] No Alert

[ ] Regular Alert (Alert Sound)

[ ] Gradual (Ambient) Alert
24. During which alert do you feel you were the second most productive? *

Mark only one oval.

□ No Alert
□ Regular Alert (Alert Sound)
□ Gradual (Ambient) Alert

25. During which alert do you feel you were the least productive? *

Mark only one oval.

□ No Alert
□ Regular Alert (Alert Sound)
□ Gradual (Ambient) Alert

26. Of the three alerts, which one was your favorite? Why? *

........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................

27. What would you use the ambient audio effect for? *

........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................

28. Do you have any other feedback for researchers about the experiment?

........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................

Thank You!
29. If you would like to keep informed about this research, please provide your email address below and we will email you any published results. (Otherwise, leave it blank and you will no longer hear from us!)
Experiment Part 2
SuperHuman Computing Lab - Audio Alert Study

* Required

1. User ID *
   (Provided by Researchers)

Audio Segment 1

2. What effect did you hear? *
   Mark only one oval.
   - Signal Distortion
   - Lowpass Filter
   - Echo Effect
   - Tremolo
   - No Effect

3. How noticeable is this effect? (Only answer if you heard an effect.)
   Mark only one oval.

   1  2  3  4  5
   Not at all   Very noticeable

4. How annoying is this effect? (Only answer if you heard an effect.)
   Mark only one oval.

   1  2  3  4  5
   Not at all   Very annoying
5. **How did the effect alter your experience listening to the music?** (Only answer if you heard an effect.)
   
   *Mark only one oval.*
   
   - The effect made the audio clip significantly worse
   - The effect made the audio clip slightly worse
   - The effect made the audio clip neither better or worse
   - The effect made the audio clip slightly better
   - The effect made the audio clip significantly better

6. **Which statement best describes your familiarity with the audio segment?**
   
   *Mark only one oval.*
   
   - I've never heard it before
   - It sounds familiar and I have maybe heard it at some point, but I do not know it well
   - I've heard it a few times and could sing/hum along to parts of it
   - I have heard it many times and I could sing/hum along to most of it from memory

7. **Optional: If you made any mistakes (e.g. clicking the button too early) describe what you meant to do here.**

   ...........................................................
   ...........................................................
   ...........................................................
   ...........................................................
   ...........................................................

**Audio Segment 2**

8. **What effect did you hear?**
   
   *Mark only one oval.*
   
   - Signal Distortion
   - Lowpass Filter
   - Echo Effect
   - Tremolo
   - No Effect

9. **How noticeable is this effect?** (Only answer if you heard an effect.)
   
   *Mark only one oval.*
   
   
   
   
   
   
   1 2 3 4 5
   
<table>
<thead>
<tr>
<th>Not at all</th>
<th>Very noticeable</th>
</tr>
</thead>
</table>
10. How annoying is this effect? (Only answer if you heard an effect.)
   Mark only one oval.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very annoying</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

11. How did the effect alter your experience listening to the music? (Only answer if you heard an effect.)
   Mark only one oval.
   - The effect made the audio clip significantly worse
   - The effect made the audio clip slightly worse
   - The effect made the audio clip neither better or worse
   - The effect made the audio clip slightly better
   - The effect made the audio clip significantly better

12. Which statement best describes your familiarity with the audio segment? *
   Mark only one oval.
   - I've never heard it before
   - It sounds familiar and I have maybe heard it at some point, but I do not know it well
   - I've heard it a few times and could sing/hum along to parts of it
   - I have heard it many times and I could sing/hum along to most of it from memory

13. Optional: If you made any mistakes (e.g. clicking the button too early) describe what you meant to do here.

   ........................................................................................................................................
   ........................................................................................................................................
   ........................................................................................................................................
   ........................................................................................................................................
   ........................................................................................................................................

Audio Segment 3

14. What effect did you hear? *
   Mark only one oval.
   - Signal Distortion
   - Lowpass Filter
   - Echo Effect
   - Tremolo
   - No Effect
15. **How noticeable is this effect? (Only answer if you heard an effect.)**

*Mark only one oval.*

1 2 3 4 5

Not at all  □  □  □  □  □  Very noticeable

16. **How annoying is this effect? (Only answer if you heard an effect.)**

*Mark only one oval.*

1 2 3 4 5

Not at all  □  □  □  □  □  Very annoying

17. **How did the effect alter your experience listening to the music? (Only answer if you heard an effect.)**

*Mark only one oval.*

☐ The effect made the audio clip significantly worse
☐ The effect made the audio clip slightly worse
☐ The effect made the audio clip neither better or worse
☐ The effect made the audio clip slightly better
☐ The effect made the audio clip significantly better

18. **Which statement best describes your familiarity with the audio segment?** *

*Mark only one oval.*

☐ I've never heard it before
☐ It sounds familiar and I have maybe heard it at some point, but I do not know it well
☐ I've heard it a few times and could sing/hum along to parts of it
☐ I have heard it many times and I could sing/hum along to most of it from memory

19. **Optional: If you made any mistakes (e.g. clicking the button too early) describe what you meant to do here.**

........................................................................................................................................................................
........................................................................................................................................................................
........................................................................................................................................................................
........................................................................................................................................................................
........................................................................................................................................................................
........................................................................................................................................................................

**Audio Segment 4**
20. **What effect did you hear? * Mark only one oval.**
- Signal Distortion
- Lowpass Filter
- Echo Effect
- Tremolo
- No Effect

21. **How noticeable is this effect? (Only answer if you heard an effect.) Mark only one oval.**
- 1
- 2
- 3
- 4
- 5

<table>
<thead>
<tr>
<th>Not at all</th>
<th>Very noticeable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

22. **How annoying is this effect? (Only answer if you heard an effect.) Mark only one oval.**
- 1
- 2
- 3
- 4
- 5

<table>
<thead>
<tr>
<th>Not at all</th>
<th>Very annoying</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

23. **How did the effect alter your experience listening to the music? (Only answer if you heard an effect.) Mark only one oval.**
- The effect made the audio clip significantly worse
- The effect made the audio clip slightly worse
- The effect made the audio clip neither better or worse
- The effect made the audio clip slightly better
- The effect made the audio clip significantly better

24. **Which statement best describes your familiarity with the audio segment? * Mark only one oval.**
- I've never heard it before
- It sounds familiar and I have maybe heard it at some point, but I do not know it well
- I've heard it a few times and could sing/hum along to parts of it
- I have heard it many times and I could sing/hum along to most of it from memory
25. Optional: If you made any mistakes (e.g. clicking the button too early) describe what you meant to do here.

..........................................................................................................................
..........................................................................................................................
..........................................................................................................................
..........................................................................................................................
..........................................................................................................................

Audio Segment 5

26. What effect did you hear? *
   
   Mark only one oval.
   
   □ Signal Distortion
   □ Lowpass Filter
   □ Echo Effect
   □ Tremolo
   □ No Effect

27. How noticeable is this effect? (Only answer if you heard an effect.)
   
   Mark only one oval.
   
   1 2 3 4 5
   Not at all □ □ □ □ □ Very noticeable

28. How annoying is this effect? (Only answer if you heard an effect.)
   
   Mark only one oval.
   
   1 2 3 4 5
   Not at all □ □ □ □ □ Very annoying

29. How did the effect alter your experience listening to the music? (Only answer if you heard an effect.)
   
   Mark only one oval.
   
   □ The effect made the audio clip significantly worse
   □ The effect made the audio clip slightly worse
   □ The effect made the audio clip neither better or worse
   □ The effect made the audio clip slightly better
   □ The effect made the audio clip significantly better
30. **Which statement best describes your familiarity with the audio segment?** *

   *Mark only one oval.*

   - I've never heard it before
   - It sounds familiar and I have maybe heard it at some point, but I do not know it well
   - I've heard it a few times and could sing/hum along to parts of it
   - I have heard it many times and I could sing/hum along to most of it from memory

31. **Optional: If you made any mistakes (e.g. clicking the button too early) describe what you meant to do here.**

   ..................................................................................................................
   ..................................................................................................................
   ..................................................................................................................
   ..................................................................................................................
   ..................................................................................................................
   ..................................................................................................................

**Audio Segment 6**

32. **What effect did you hear?** *

   *Mark only one oval.*

   - Signal Distortion
   - Lowpass Filter
   - Echo Effect
   - Tremolo
   - No Effect

33. **How noticeable is this effect? (Only answer if you heard an effect.)**

   *Mark only one oval.*

   1 2 3 4 5

   Not at all   Very noticeable

34. **How annoying is this effect? (Only answer if you heard an effect.)**

   *Mark only one oval.*

   1 2 3 4 5

   Not at all   Very annoying
35. **How did the effect alter your experience listening to the music?** (Only answer if you heard an effect.)
   
   *Mark only one oval.*
   - The effect made the audio clip significantly worse
   - The effect made the audio clip slightly worse
   - The effect made the audio clip neither better or worse
   - The effect made the audio clip slightly better
   - The effect made the audio clip significantly better

36. **Which statement best describes your familiarity with the audio segment?** *
   
   *Mark only one oval.*
   - I've never heard it before
   - It sounds familiar and I have maybe heard it at some point, but I do not know it well
   - I've heard it a few times and could sing/hum along to parts of it
   - I have heard it many times and I could sing/hum along to most of it from memory

37. **Optional: If you made any mistakes (e.g. clicking the button too early) describe what you meant to do here.**

   ........................................................................................................................................
   ........................................................................................................................................
   ........................................................................................................................................
   ........................................................................................................................................
   ........................................................................................................................................
   ........................................................................................................................................

**Audio Segment 7**

38. **What effect did you hear?** *
   
   *Mark only one oval.*
   - Signal Distortion
   - Lowpass Filter
   - Echo Effect
   - Tremolo
   - No Effect

39. **How noticeable is this effect?** (Only answer if you heard an effect.)
   
   *Mark only one oval.*

   1 2 3 4 5
   Not at all  .................................................................................................................................. Very noticeable
40. How annoying is this effect? (Only answer if you heard an effect.)
   Mark only one oval.

   1  2  3  4  5

   Not at all  ☐  ☐  ☐  ☐  ☐  Very annoying

41. How did the effect alter your experience listening to the music? (Only answer if you heard an effect.)
   Mark only one oval.

   ☐ The effect made the audio clip significantly worse
   ☐ The effect made the audio clip slightly worse
   ☐ The effect made the audio clip neither better or worse
   ☐ The effect made the audio clip slightly better
   ☐ The effect made the audio clip significantly better

42. Which statement best describes your familiarity with the audio segment? *
   Mark only one oval.

   ☐ I've never heard it before
   ☐ It sounds familiar and I have maybe heard it at some point, but I do not know it well
   ☐ I've heard it a few times and could sing/hum along to parts of it
   ☐ I have heard it many times and I could sing/hum along to most of it from memory

43. Optional: If you made any mistakes (e.g. clicking the button too early) describe what you meant to do here.

   ....................................................................................................................
   ....................................................................................................................
   ....................................................................................................................
   ....................................................................................................................
   ....................................................................................................................

Audio Segment 8

44. What effect did you hear? *
   Mark only one oval.

   ☐ Signal Distortion
   ☐ Lowpass Filter
   ☐ Echo Effect
   ☐ Tremolo
   ☐ No Effect
45. How noticeable is this effect? (Only answer if you heard an effect.)
   Mark only one oval.

   1  2  3  4  5

   Not at all  ○  ○  ○  ○  ○  Very noticeable

46. How annoying is this effect? (Only answer if you heard an effect.)
   Mark only one oval.

   1  2  3  4  5

   Not at all  ○  ○  ○  ○  ○  Very annoying

47. How did the effect alter your experience listening to the music? (Only answer if you heard an effect.)
   Mark only one oval.

   ○  The effect made the audio clip significantly worse
   ○  The effect made the audio clip slightly worse
   ○  The effect made the audio clip neither better or worse
   ○  The effect made the audio clip slightly better
   ○  The effect made the audio clip significantly better

48. Which statement best describes your familiarity with the audio segment? *
   Mark only one oval.

   ○  I've never heard it before
   ○  It sounds familiar and I have maybe heard it at some point, but I do not know it well
   ○  I've heard it a few times and could sing/hum along to parts of it
   ○  I have heard it many times and I could sing/hum along to most of it from memory

49. Optional: If you made any mistakes (e.g. clicking the button too early) describe what you meant to do here.

   ........................................................................................................
   ........................................................................................................
   ........................................................................................................
   ........................................................................................................
   ........................................................................................................

Audio Segment 9
50. **What effect did you hear?** *

   *Mark only one oval.

   - Signal Distortion
   - Lowpass Filter
   - Echo Effect
   - Tremolo
   - No Effect

51. **How noticeable is this effect?** (Only answer if you heard an effect.)

   *Mark only one oval.

   1  2  3  4  5

   Not at all   | Very noticeable

52. **How annoying is this effect?** (Only answer if you heard an effect.)

   *Mark only one oval.

   1  2  3  4  5

   Not at all   | Very annoying

53. **How did the effect alter your experience listening to the music?** (Only answer if you heard an effect.)

   *Mark only one oval.

   - The effect made the audio clip significantly worse
   - The effect made the audio clip slightly worse
   - The effect made the audio clip neither better or worse
   - The effect made the audio clip slightly better
   - The effect made the audio clip significantly better

54. **Which statement best describes your familiarity with the audio segment?** *

   *Mark only one oval.

   - I've never heard it before
   - It sounds familiar and I have maybe heard it at some point, but I do not know it well
   - I've heard it a few times and could sing/hum along to parts of it
   - I have heard it many times and I could sing/hum along to most of it from memory
55. Optional: If you made any mistakes (e.g. clicking the button too early) describe what you meant to do here.

........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................

Audio Segment 10

56. **What effect did you hear?** *
   
   *Mark only one oval.*
   
   - Signal Distortion
   - Lowpass Filter
   - Echo Effect
   - Tremolo
   - No Effect

57. **How noticeable is this effect?** (Only answer if you heard an effect.)
   
   *Mark only one oval.*

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

58. **How annoying is this effect?** (Only answer if you heard an effect.)
   
   *Mark only one oval.*

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

59. **How did the effect alter your experience listening to the music?** (Only answer if you heard an effect.)
   
   *Mark only one oval.*

   - The effect made the audio clip significantly worse
   - The effect made the audio clip slightly worse
   - The effect made the audio clip neither better or worse
   - The effect made the audio clip slightly better
   - The effect made the audio clip significantly better
60. **Which statement best describes your familiarity with the audio segment?** *

*Mark only one oval.*

- I've never heard it before
- It sounds familiar and I have maybe heard it at some point, but I do not know it well
- I've heard it a few times and could sing/hum along to parts of it
- I have heard it many times and I could sing/hum along to most of it from memory

61. **Optional: If you made any mistakes (e.g. clicking the button too early) describe what you meant to do here.**

..........................................................................................................................................................
..........................................................................................................................................................
..........................................................................................................................................................
..........................................................................................................................................................
..........................................................................................................................................................
..........................................................................................................................................................

**All Done!**

62. **Optional: Do you have any other feedback about how the experiment went today?**

..........................................................................................................................................................
..........................................................................................................................................................
..........................................................................................................................................................
..........................................................................................................................................................
..........................................................................................................................................................
..........................................................................................................................................................

---

SUPERHUMAN COMPUTING LAB
at the University of Colorado Boulder