Data Sonification with Natural Sounds

MD NAIMUL HOQUE, University of Maryland, College Park
MD EHTESHAM-UL-HAQUE, Pennsylvania State University
NIKLAS ELMQVIST, University of Maryland, College Park, USA
SYED MASUM BILLAH, Pennsylvania State University

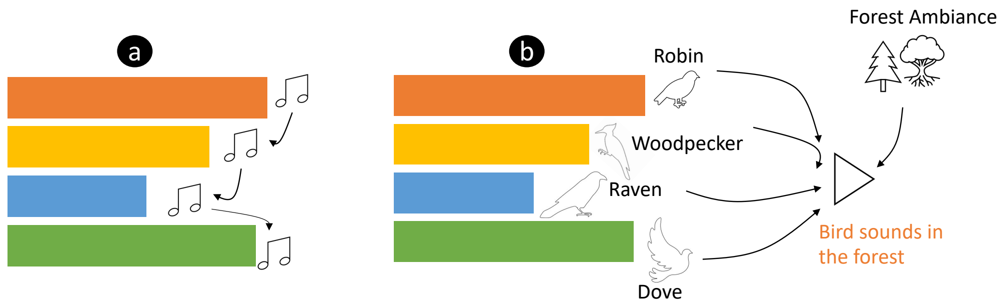


Fig. 1. An illustration of sonifying a horizontal bar chart with existing approaches vs. Ambient-Sonic. (a) An existing sonification tool, Highcharts [13], plays a musical note four times *serially* to sonify four bars, where the frequency or pitch of a note is proportionate to the value (i.e., width) of that bar. (b) In contrast, Ambient-Sonic plays four ambient sounds *parallelly* to sonify four bars all at once. Each ambient sound is drawn from an ambient theme (e.g., forest and birds), and the volume of a sound is proportionate to the value (i.e., width) of that bar. For instance, we have drawn four bird sounds (e.g., Robin, Woodpecker, Raven, and Dove) to map with four bars and use a forest ambiance as background, thus sonifying the bar chart as listening to bird sounds in the forest.

Ambient natural sounds, such as the sound of raindrops, moving water, forest ambiance, and birdsong, have positive impacts on humans. These sounds are used in mindfulness and meditation to calm the mind, and in therapy to reduce pain, lower stress, and enhance moods. Each natural sound has a distinctive pattern that humans can recognize from a mix of multiple natural sounds. In this paper, we propose a sonification method for translating visual representation of data, such as a bar chart, line chart, or scatter plot to a mix of natural sounds. To support our proposed method, we designed a sonification prototype, Ambient-Sonic. The prototype is designed in two phases. First, we conducted a pilot study with 18 participants to asses the potential of representing data with natural sounds. The study revealed that it is possible to sonify multiple categories of a nominal data parallelly using natural sounds, allowing a user to compare categories easily. This mechanism is different than the existing sonification techniques where categories are typically played serially, one after one, requiring users to memorize the categories for performing effective comparison between them. Informed by the study, we identified design guidelines for mapping data to natural sounds and refined our initial version of the tool. Finally, we conducted a user study with 14 participants and found that Ambient-Sonic increased participants' efficiency in understanding the trend of a dataset and inferring summary statistics from a dataset, in comparison to existing sonification tools.

CCS Concepts: • **Human-centered computing** → **Visual analytics**.

Authors' addresses: Md Naimul Hoque, University of Maryland, College Park, nhoque@umd.edu; Md Ehtesham-Ul-Haque, Pennsylvania State University, mfe5232@psu.edu; Niklas Elmqvist, elm@umd.edu, University of Maryland, College Park, College Park, MD, USA; Syed Masum Billah, Pennsylvania State University, sbillah@psu.edu.

2021. Manuscript submitted to ACM

1 INTRODUCTION

Data sonification is analogous to data visualization, where one perceives data by engaging auditory sense instead of visual perception. In technical terms, sonification is defined as the mapping of abstract data to non-speech sound [38]. Such auditory representation can complement information visualization and create a multi-modal interaction for perceiving complex data [54]. Researchers have used sonification to effectively perceive complex data, such as protein structure [51, 52], cosmic objects [15, 31], and charts [16, 45].

In this paper, we propose to use ambient natural sounds (e.g., birds chirping) as alternatives to musical notes, typically used in data sonification. Our work is motivated by three observations. *First*, humans are familiar with natural sounds by birth. We perceive these sounds on a daily basis while walking, running, or taking a deep breath in a green mountain top. Noticeably, we can distinguish these sounds individually even when multiple natural sounds are playing parallelly. Based on this observation, we hypothesize that multiple categories of nominal data can be sonified together where each category can be mapped to a different natural sound, allowing a user to compare the categories easily and obtain an overview of the data. In contrast, existing sonification tools sonify each category serially, one after one [1, 13]. While this mechanism has seen success previously, a user needs to remember the value of a category to compare it with other categories since at a time only one category is sonified.

Second, in musical contexts, mapping data to higher-order musical dimensions, such as tempo, form, and timbre, requires musically trained ears to realize subtle changes in data [38]. In contrast, humans are naturally trained to listen to natural sounds. Thus, a natural sound based sonification tool may reduce training needed to understand data from audio and make it less overwhelming for people who do not necessarily possess musically trained ears.

Finally, an additional benefit of listening to a mix of natural sounds is that it has positive impact on health and wellbeing [12], including calming the mind [36], reducing stresses [5, 34], increasing attention [10, 34], and enhancing mood [24, 25]. Moreover, these sounds evoke a similar response in the human brain's early visual cortex [44], regardless of a person's visual ability.

To assess the validity of our proposed method, we at first conducted a pilot study with 18 participants. We developed an initial prototype for the study, Ambient-Sonic-Lite, that can sonify bar charts and line charts. This study revealed that participants could accurately reproduce line charts and bar charts from the mix of natural sounds. However, the choice of sound sources and the number of data categories impacted their perception.

We drew insights from these findings and identified six guidelines for mapping data to natural sounds. We refined the initial prototype accordingly and conducted another study with 14 participants to compare the refined prototype, Ambient-Sonic, with existing musical note-based sonification tools. The second study revealed that Ambient-Sonic improved participants' performance in understanding trends and inferring summary statistics such as *min* and *max* from datasets by 12.49%. Further, the improvements were 14.11% for bar charts and 8.58% for line charts.

In sum, our contributions are as follows:

- We show that natural sounds have distinct expressive patterns that can be harnessed to represent data.
- We propose Ambient-Sonic, a sonification prototype that uses ambient natural sounds to sonifiy common chart types (e.g., Bar, Scatter, and Line).
- We outline a list of sonification guidelines for mapping data to natural sounds, identified from a study with 18 participants.
- We report on a user study with 14 participants to evaluate the effectiveness of Ambient-Sonic in representing summary statistics and trends in data compared to existing sonification tools.

We make no claim that our method is a one-size solution that will fit all users. Rather, we see Ambient-Sonic as a way to complement existing sonification techniques with natural sounds. We conclude this paper with a discussion on ways to merge research using musical notes and natural sounds and how together they can advance sonification research in Section 7.

2 BACKGROUND AND RELATED WORK

Our work is related to prior efforts on alternate and multi-modal data representation, especially with sonification. We discuss research on these areas below.

2.1 Alternate and Multi-modal Data Representation

While data visualization remains the prominent way for data analysis, researchers have explored alternate sensory medium such as *audio, touch, and smell* for representing data. Examples of audible charts include Apple Audio Graph [1, 2], Highcharts [3, 13], Sonification Sandbox [14, 46], TwoTone [4], Vega-lite [37], and iSonic [54]. Sonification has also seen success in scientific endeavors such as analysis of protein structure [51, 52], heart rate variability [6], and cosmic objects [31]. Other applications such as accessible floor plans [20] and natural language description of line charts [17] have also appeared. These tools use musical notes and pitches to convey data. We, on the other hand use easy-to-understand natural sounds for data representation. We believe the simultaneous auditory feedback generated by multiple natural sounds will increase the flux of information conveyed and allow users to perceive data efficiently without any significant training.

Touch senses and tactile interfaces are also common for representing data. Existing works include physical bar charts [16, 18, 42, 43], tactile feedback based interpretation of 2D information [50], 3D printed tangible maps [22, 23], wheeled robots for physical data visualizations [28], and others. Recently, Patnaik et al. [32] proposed an olfactory system to convey information through scents. A follow-up paper based on this work provided a ranking of the sensory channels through a perception study [9]. While effective, touch and smell-based systems often require external hardware and are only applicable to specific types of data (e.g., bar charts), making their applicability limited to a lab or to a data domain.

We note here that several of the above mentioned works have used data sonification and other alternate representation to increase data accessibility. While we take inspiration from these works, our focus in this paper was to understand the principles and potentials of natural sounds for sonification purpose. For this initial stage, we strictly evaluated our method as a sonification technique, not as an accessibility tool, especially from the perspective of sighted users. We do however believe that our research will inform any future data accessibility methods that might use data sonification. We discussed these implications in Section 7. The rest of this section is designed to provide a detail overview of the state-of-the-art sonification tools and techniques.

2.2 Taxonomy of Sonification

Sonification can be divided into different categories depending on the function or technique of the task [21]. Function-based sonification is of four types - alert and notification, status indicator, data exploration, and art, entertainment, sports, and exercise. Based on the technique, sonification has three categories - event-based/parameter mapping sonification(PMSon), model-based, and continuous sonification. A particular sonification can incorporate multiple categories. Some auditory displays, such as soundscapes [29], blend status indicator and data exploration functions.

Data exploration may employ event-based approaches, model-based sonification, or continuous sonification depending upon the specific task of the user [7].

2.3 Exploration Tasks for Quantitative Data

Researchers have described different types of data exploration tasks for quantitative data [19, 30, 33]. Among these types, 5 are the most common - Point Estimation, Point Comparison, Trend Identification, Identification of Data Structure, and Exploratory Inspection [21]. Point Estimation requires focusing on a single data point and finding out the detail of the data point. In sonification, a user listens to the data to find the point of interest and estimates the magnitude of the data represented by the sonification. For this task, the user may need to compare the magnitude of the data to a baseline tone. Point comparison is performed by estimating multiple data points and comparing the magnitudes. However, prior research in sonification did not contribute much to this task. Trend identification is another task that requires understanding the overall pattern of the data. Depending on the data, a user might be interested in both global and local trends. Similar to trend identifications, users might be interested in understanding the overall structure of the data to identify complex relations among multiple relations. Exploratory inspection is performed by inspecting the data without any particular tasks. This is a viable sonification task because of the temporal resolution of the auditory system and pattern detection acuity. With an exploratory analysis of sonification, users might be able to find patterns or anomalies in data, which might not be apparent from a visual representation of the data.

Zhao et. al. [53] adopted the methods of visual information searching for Parameter Mapping Sonification, and introduced four principles known as Auditory Information Seeking Principles (AISP). These principles are - (i) gist: the capability of hearing an essential summary of data in a given region, (ii) navigate: the ability to scan sequential orderings of data, (iii) filter: the capability to selectively seek data according to specific criteria, and (iv) details on demand: the ability to obtain details of one or more data collections for comparison.

Ambient-Sonic currently supports all 5 exploration tasks discussed above. Using Ambient-Sonic, a user can listen to multiple data points or categories parallelly. This mechanism helps in identification of trend, and data structures and understanding the gist of the data. We further provide easy keyboard shortcuts for interacting with the sonification. The interactions help a user navigate, filter, and obtain details on demand which aid the process of point estimation and comparison tasks.

2.4 Data to Sound Mapping

2.4.1 Sound Properties. Mapping data to appropriate sound properties is one of the challenging tasks in sonification. Non-visual senses (e.g., hearing) do not have the same capacity as sight, which is known as the bandwidth problem. Because of this restriction, sound cannot be used to provide the same amount of information compared to visualization. To address this problem, a sonification needs to - (i) maximize the amount of information carried in the sounds and (ii) reduce the amount of information presented (i.e. filter information) [21].

Researchers have proposed different audio properties for mapping different data types. Moreover, different sound dimensions are perceived differently by the listeners. For example, timbre is perceived as a categorical property whereas frequency and intensity are perceived as continuous properties [21]. Walker [47] found the pitch of the sound to be effective in mapping temperature rather than tempo. Barrass [7, 8] suggested using categorical properties such as timbre to represent nominal/categorical data types. Interval data may be represented by more continuous acoustic variables, such as pitch or loudness [48].

2.4.2 Polarity. Mapping the polarity of a sound property to data is not always straightforward because listeners might interpret it differently. Walker [47] represented temperature with the pitch of the sound. He found that some participants assumed increasing pitch represents increasing temperature (a positive polarity), while others interpreted increasing pitch to be associated with decreasing temperature (a negative polarity). He listed the preferred polarities for different mappings of data to sound and reported that if the polarity of the mapping does not match with the listeners' preferences, it impacts their performance [47, 48].

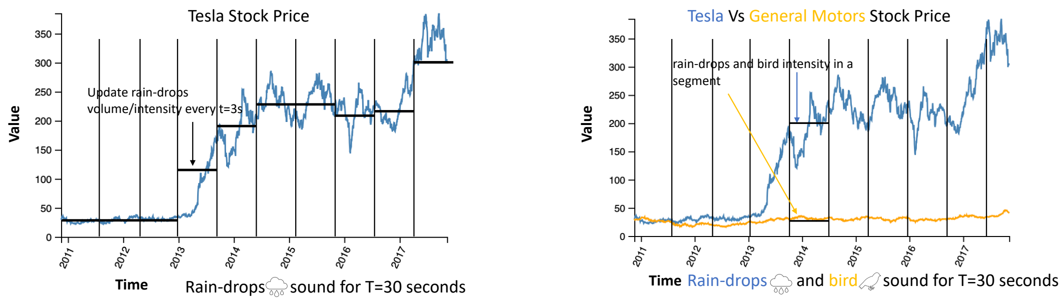
2.4.3 Scale. Walker [47, 48] empirically identified scaling factors while mapping temperature data ranging from 0 to 30 degrees of Celsius using the frequency of sound. One choice for this mapping is to scale the data to fill the entire hearing range for the human ear (20 Hz to 20000 Hz). However, the human ear is more sensitive to frequencies ranging from 1000 Hz to 5000 Hz, and mapping the data in this range was found to be more effective. Mapping data on a musical scale (e.g., notes on a piano) has been employed by different sonification tools [49]. However, such scales have a limited granularity forcing the designer to scale the entire data with only a few musical notes (e.g., MIDI notes from 35 to 100 [11]). A designer can round the data to fit the scale or employ pitch bending to play the data at the exact pitch, which can take away from the intended musicality of the approach. Popular sonification tools such as Sonification Sandbox [49] provide rounding and exact scaling option to facilitate the appropriate mapping for a particular task [21].

2.4.4 Context. Smith and Walker [40] found adding context cues can provide useful information to a sonification. For example, adding a series of clicks can help the listener keep track of the time better. They showed that when the clicks played at twice the rate of the sounds representing the data can have the effect of the major and minor tick marks on the x-axis of a visual graph. Providing a repeating reference tone representing the maximum value of the data set improved the estimation of exact data values in a data exploration task, whereas representing the starting value did not have such improvement [21].

2.5 Natural Sound for Sonification

Sonification of data with natural ambient sound has several advantages over musical sounds. Kilander and Lönnqvist [26] reported that monitoring mechanical activities (e.g., network or server performance) for a long time can be monotonous. To tackle this problem, they introduced the notion of a Weakly Intrusive Ambient Soundscape (WISP), which states that sound cues for environmental and process data should be subtle and minimally intrusive. They emphasized that anticipated sounds tend to slip from the attention of the user, and helps achieve weak intrusion. For example, ambient sounds such as a ticking clock are readily perceived and attended to when it is introduced into the environment. However, continuous ticking reaches a steady state after some period and does not create any intrusion unless the speed, timbre, or intensity of the clock tick changes to attract the attention of the listener. Kilander and Lönnqvist [26] also suggested using and modulating sounds that fit well with the acoustic ecology of the process monitor's environment to achieve minimum intrusiveness. To increase the quality of the acoustic ecology further, they used real-world sounds (easily recognizable and natural) rather than synthesized noises and musical tones. Stockman described how blind people can make use of environmental sounds, and also discussed how these sounds can be supplement with technology-generated sounds for different tasks [41]

Language and musical sounds tend to be culture-specific, which can be difficult for listeners to find the relation between the sounds and their meaning. On the contrary, the semantics of natural ambient sound is more universal. Sonification techniques relying on sounds that are encountered commonly from the environment are likely to be more culturally independent [21].



(a) Tesla's stock price represented using a T=30 seconds long rain sound. To depict the overall trend in the time series, Ambient-Sonic-Lite updates the volume of the sound every t=3s proportionate to the average value of the segment.

(b) Comparison between Tesla and General Motor's stock trend. Tesla is represented using the same rain drop sound whereas general motors is represented using a bird chirping sound.

Fig. 2. Ambient-Sonic-Lite transforming time series data to natural sounds.

3 FEASIBILITY STUDY

To assess the potential of natural sounds for sonification, we conducted a feasibility study with 18 participants. We particularly aimed at finding potential limitations of our approach to identify guidelines for mapping data to natural sounds.

3.1 Participants

The participants were recruited through local mailing lists, university mailing lists, and social media posts. Our inclusion criteria included familiarity with basic data visualizations, owning a stereo headphone, and having no hearing impairments. Participants varied in age from 19 to 35 ($M = 25$, $SD = 4.21$) and gender ($male = 9$, $female = 9$), and were either major in STEM or pursuing a STEM degree. None of the participants reported prior experience with data sonification. Four participants reported prior musical training, either as a self-taught musician, or having formal training. All participants reported familiarity with basic data visualization such as a bar, line, or scatter plot. Participation was voluntary with no compensation. We referred to the participants as P1-P18 below.

3.2 Apparatus

For this study, we took inspiration from the existing sonification tools and developed a natural sound-based sonification tool: *Ambient-Sonic-Lite*. Ambient-Sonic-Lite was web-based and utilized the browser-independent *Web Audio Api* [1]. We collected the following list of freely available ambient sound sources from the web: rain, crickets, frogs, rain, owls, chimes, distant rolling thunder, woodpeckers, nightingale, dry leaves, owl, cuckoo, wood-burning, grasshopper, wind howling, shore waves, soft wind, and seagull. We implemented two types of data visualization in this initial version: (1) line chart and (2) bar chart. We further included multi-line charts to assess the possibility of mixing multiple natural sounds together.

In a line chart, the height of the line represents its value at a point. Existing systems utilize the pitch or frequency of digitally created musical notes to sonify a line chart [1, 13]. However, natural sound sources are not digital in nature.

[1]https://developer.mozilla.org/en-US/docs/Web/API/Web_Audio_API

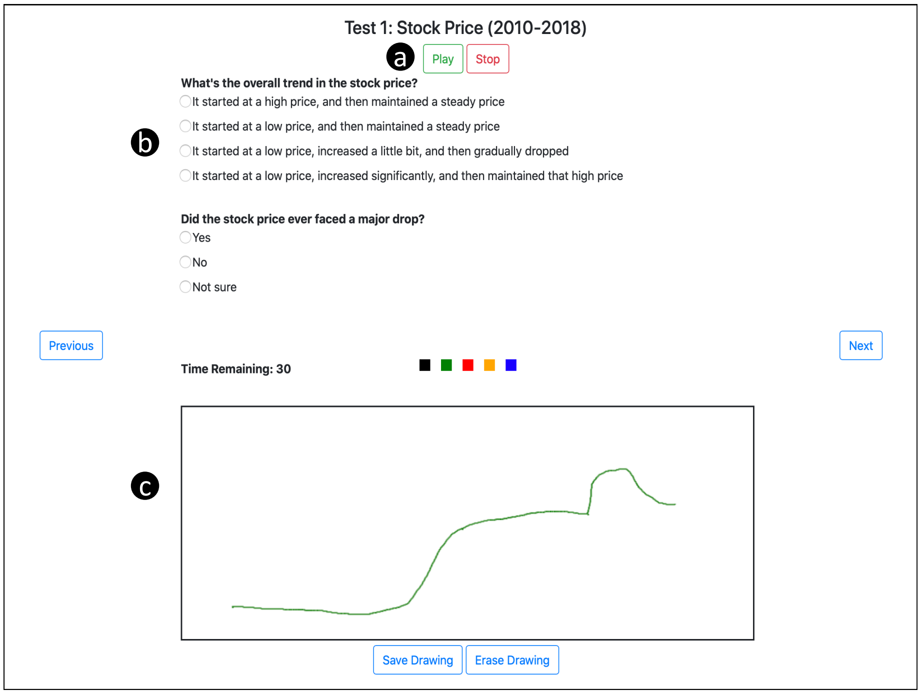


Fig. 3. Web interface for the pilot study. (a) Buttons (Play, Stop) to control the audio; (b) Multiple choice Q&A; and (c) A canvas to sketch the inferred patterns from the audio.

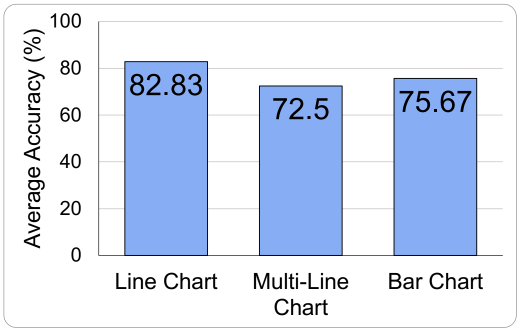
They already have their own frequency distribution. Thus, instead of using frequency to sonify a line chart, we used the volume of the sound sources for this purpose. Ambient-Sonic-Lite sonifies a line chart, or a multi-line chart through a T seconds long audio consisting of natural sounds. We periodically change the volume of the sound sources to indicate the trend. In the case of Figure 2a, the volume of the rain-drops (Tesla) increases gradually over time to follow the stock trend. Based on this principle, Ambient-Sonic-Lite can also sonify multiple time series (Figure 2b).

In a bar chart, a user should be able to identify the number of bars as well as their relative values for comparison. To support that, existing tools sonify the bars serially (left to right), one after one. We followed this method to play the bars serially. For example, if a bar chart has five bars, a user would first listen to a beep sound and then a natural sound (e.g., raindrop) with its volume set in proportional to the bar's height. The same process is carried out for the other four bars in a linear fashion.

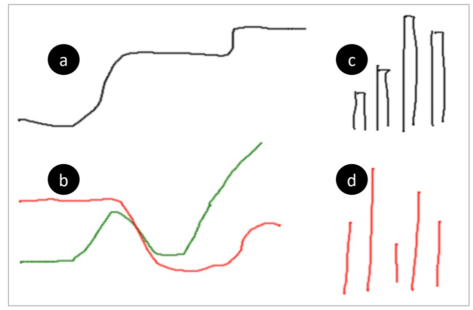
3.3 Method

We deployed Ambient-Sonic-Lite on a public web server and shared its URL with the participants. Each session was administered by one of the authors over teleconferencing software (e.g., Skype or ZOOM). After consent, participants shared their screens. The administrator communicated with the participants, providing verbal instructions on how to perform each task. Each session was screen capture and transcribed by one of the authors.

A session consisted of three parts, reflecting three visualizations supported by Ambient-Sonic-Lite : line, multi-line, and bar. For each part, we followed the protocol described below: Each participant listened to two training audio clips



(a) Average accuracy in answering multiple-choice questions.



(b) Sample sketches from the study.

Fig. 4. Pilot study results.

along with the visualization of the input dataset. Participants could listen to these clips as many times as they wished. Once they became comfortable with Ambient-Sonic-Lite, we provided them with three clips (one after one) without showing the visualizations. Each clip contained a mix of natural sounds representing a hidden input dataset. They were asked to listen to a clip as many times as they wished and then answer several multiple-choice questions designed to measure their comprehension of data from the audio clip. We also asked them to sketch a visualization of the input dataset to measure their auditory perception of data. After completing all three clips in a part, participants moved to the next part and followed the same protocol. We maintained the same task order for all participants: line charts, followed by multi-line charts, and ended with bar charts, because the latter tasks depended on the interpretation of the former ones. After the question answering part, we revisited the questions with the participants where participants could see the hidden visualization and their answers to the questions. This part was designed to understand the limitations of our tool and how we could address them in a future version.

Figure 3 presents the web interface used for the study. Participants used a canvas (Figure 3c) to sketch a visualization of the original input data. The multiple-choice questions reflected common data analysis tasks, such as finding *min*, *max*, and *average*; identifying the total number of categories; and comparing data points. These questions also reflected the common sonification tasks discussed in Section 2.3.

3.4 Findings

Figure 4a presents participants' average accuracy in answering multiple-choice questions. Participants were able to answer inference questions with an average accuracy of 82.83% for line charts. However, the average accuracies were 72.5% for multi-line charts and 75.67% for bar charts. We further provide a critical assessment of users' feedback below.

3.4.1 Blending Audio Sources. In the case of multi-line charts, multiple sound sources such as rain, ocean waves, bird, and cricket were used to represent multiple lines. Although participants understood the presence of different sound sources easily, understanding the trend from the sounds was difficult for them once the number of sound sources (n) was more than 3 (P1-3, P5, P8, P15-16). Participants suggested that the choices for the individual sound sources are critical in these scenarios. They mentioned that a few sound sources attracted their attention more than the other ones.

I think some sounds did not go along well. Some sounds were inherently louder than others. I would say the sounds need to be chosen very carefully. (P12)

The cricket sound was very timid even when it had a high volume. I would probably not pay attention to such sounds. On the other hand, the raven had a big sound, something I can imagine to be representing a high quantity. (P8)

We noticed this problem was particularly evident when white noise sounds such as rain, ocean waves, etc. were used for representing data. Participants mentioned that with loud white noises it was hard to differentiate changes in other sounds such as birds, crickets, etc.

When you use rain sound with a bird sound, the pitch of them is very different. I felt like the bird sound was sharper and I could detect the changes in volumes easily. However, it was difficult for me to compare it with the rain sound which was a completely different sound than the bird. (P4)

The bird sound had a distinct pattern that I was able to track easily. I did not find any patterns in the rain sound. It was always a wholesome sound. (P7)

Participants suggested that selecting sounds that are comparable and have the same perception of loudness would be effective (P3, P5, P8). Further, participants suggested that isolating individual sound sources may reduce complexity in such scenarios.

When there were 4 different sounds, I could only concentrate on two different sounds at a time. It would be nice to have the option to listen to the full audio and individual sound sources pair-wise. (P10)

3.4.2 Memorization. Ambient-Sonic-Lite sonified a bar chart serially (one after another). Participants understood this serial presentation easily. However, it was difficult for them to compare bars in this way (P1-P4, P6-10, P13, P16-18). We noticed their accuracy was around 50% for questions where participants were asked to compare non-adjacent bars.

I had to memorize the intensity of each bar so that I can compare. I listened to the audio multiple times to sort the values. (P17)

First of all, I needed to remember the volume of each bar. By the time I listened to the fifth bar, I already forgot the first one. (P15)

3.4.3 Changes in Volume. Finally, we divided timelines into ten equal parts and set the volume of a segment proportionate to its average value. This mechanism did not result in a frequent change in volume. When comparing the actual visualization with their own drawings in the post-study discussion, participants identified a few cases where the volume did not change appropriately to convey the trend (P2, P6-9, P14).

4 MAPPING DATA TO NATURAL SOUNDS

Based on the findings from the study (Section 3.4), we identified six guidelines for mapping data to natural sounds. We describe the refined sonification mapping below. We link the findings from the study whenever applicable.

G1. Separate white noises. One of the findings from the study suggests that white noises may not be ideal for mapping data as loud white noises can be distracting, irritating and can consume other sounds. To investigate further, we computed spectrogram for several sound sources. Figure 5a presents spectrogram for a rain sound, a white noise. We notice that its frequency distribution is expanded to the full spectrum with high amplitudes. In contrast, a raven and a woodpecker only take a fixed portion of the frequency spectrum. Thus, a white noise with high volume could consume

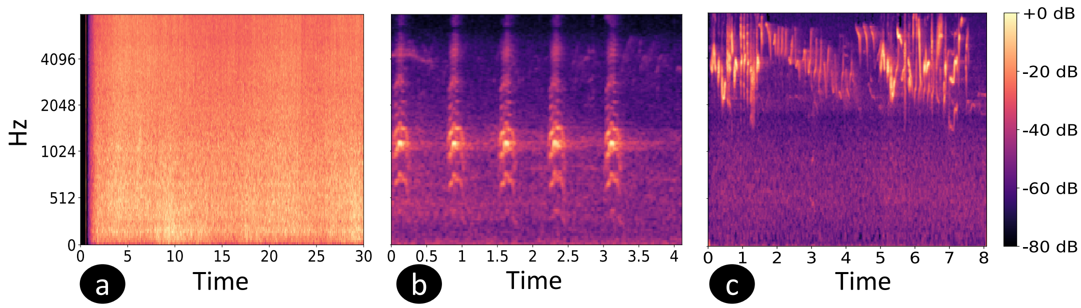


Fig. 5. Mel Spectrogram of three sound sources: (a) rain, (b) a bird sound (Raven); and (c) a woodpecker. X-axis represents time (seconds) while Y-axis represents frequency (Hz). The color scale represents amplitude in decibel (dB).

other sounds and make it hard to listent to any other sounds. They should not be used for representing data; rather they should be used as background sounds to create ambiance such as a rainy day, or a forest environment.

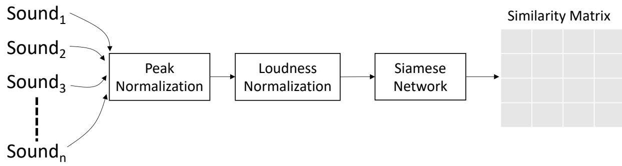
G2. Choose similar type of sounds with distinct patterns. The findings from the study suggests that the selection of individual sound sources is critical for effective sonification. While G1 will ensure that white noises will not be used for mapping data, there exists several other design constraints. For effective mapping, the sound sources need to have distinct patterns and frequencies so that they are easily distinguishable. However, the sounds should not be of completely different type (e.g., a elephant sound and a cricket sound) as that will make it harder to compare them based on volume. Thus, the sound sources need to have an underlying similarity (e.g., different birds) in them that they are comparable yet distinguishable when played together (e.g., woodpecker, raven).

G3. Normalize and calibrate loudness. The volume or loudness of the sound sources need to be normalized first to be on the same scale. Standard auditory compression and normalization can be applied to achieve this. However, our initial findings suggest that the perception of loudness may vary person to person. Thus, users should have the ability to validate the normalization performed automatically as well as be able to calibrate them to their own perception. After that, the volume of the sources can be set in proportion to the quantitative values. This will ensure that users will perceive the values correctly through the changes in volume.

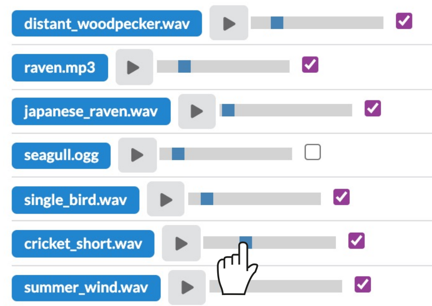
G4. Continuous changes in volume. Separating a line into a small number of segments and making discrete changes to the volume based on that can hide drastic changes in data. Thus, our refined tool should use a larger number of segments to change volume more frequently, rather discretely.

G5. Sonify categorical data simultaneously. Ambient-Sonic-Lite sonfied bars serially, one after another. However, our findings suggest that participants struggled to compare bars adequately in this way. In contrast, we observed that participants understood the presence of multiple sound sources in multi-line charts, despite some difficulties with white noises. Thus, following the similar principle of multi-line chart, we can sonify each bar with a distinct sound source and play them simultaneously, enabling a better mechanism to represent a bar chart similar to its visual counterpart.

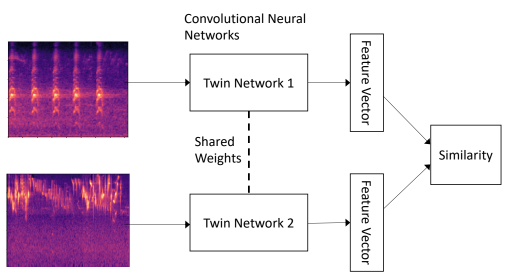
G6. Interactions for separating sound sources. Participants struggled to understand data with more than 3 categories in the pilot study. While this was partly because of the sound sources chosen and G1-3 will hopefully mitigate this concern partly, intuitive interactions can further reduce complexity of the audio and help users listen to a single data



(a) Preprocessing sound sources



(b) Calibrating volumes



(c) Simaese Model for sound similarity

Shortcut	Function
SPACE	Play or stop the sonification.
UP ARROW	Increase speed of the audio.
DOWN ARROW	Decrease speed of the audio.
1 to 9	Toggle a datapoint for sonification.
ESC	Reset data selection.

(d) Keyboard shortcuts for interaction

Fig. 6. **Components of Ambient-Sonic** (a) Normalizing loudness of different sound sources and finding similarity between them. (b) Interactive panel for selecting sound sources and calibrating their volumes. (c) Model architecture for computing similarity between sound sources. (d) Keyboard shortcuts for interactions in Ambient-Sonic.

point (a sound source) or a group of data points (multiple sound sources). However, the number of categories (parallel sound sources) should be limited to 5 since each category will inadvertently increase complexity to the overall audio.

5 AMBIENT-SONIC

Based on the design guidelines outlined in Section 4, we refined our initial prototype. In this section, we present the components of our refined sonification prototype, Ambient-Sonic.

5.1 Ambient Themes

According to G1, we at first separated the white noises from the collection of sound sources. Further, according to G2, the sound sources need to be comparable. To facilitate that, we created a collection of different bird sounds that are freely available on the web. Ambient-Sonic uses the white noises as background sounds while the bird sounds are used for representing data. Thus, together they create an ambient theme (e.g., bird sounds in the forest). It is worth noting here that although we created one theme for the purpose of demonstration, different types of sounds may create different themes (e.g., a coffee house, walking in a city) and they will be supported the same way.

5.2 Preprocessing Natural Sounds

5.2.1 Loudness Normalization. To support G2 and G3, we introduced a preprocessing pipeline in Ambient-Sonic (Figure 6a). At first, we performed two sorts of volume or loudness normalization: (1) Peak Normalization; and (2) Loudness Normalization. The purpose of the normalization is to have every sound sources on the same loudness level. Peak Normalization scans the sound sources and sets any signal over 0 dBFS to 0 dBFS. Note that 0 dBFS (Decibels relative to full scale) is considered the loudest level for an audio signal.

Loudness normalization converts average amplitude of an audio signal to a predefined reference level. For example, Youtube uses -14 LUFS (Loudness Units relative to Full Scale) as their reference level and converts every audio to that level. This ensures that sounds played one after one have the same average amplitude. We used the same method to normalize the loudness of our sound sources. We used -15 LUFS as our reference level.

5.2.2 Audio Similarity. According to G2, we need to use sounds that have distinct patterns in them. In other words, the sounds need to possess dissimilarities between them so that they are less likely to overlap in their frequency distribution. Thus, we introduced a module in our preprocessing pipeline to compute similarity between the bird sounds so that we can choose the bird sounds that are most dissimilar for sonification.

To find similarity between different sound sources, we introduced a siamese neural network [27]. A siamese network is a one-shot learning method that learns to optimize similarity between objects, instead of predicting a class for an object. Koch et al. [27] showed this objective function to have high discriminative power with only one sample per class. Since then siamese models have been widely used for face recognition, and object detection. We opted for this model as we only had a limited number of sound sources and siamese network may generalize to newly added unseen sound sources. Figure 6c presents the model. It has two branches each with a Convolutional Neural Network (CNN). We followed the architecture introduced by Koch et al. [27] for the CNNs. Both CNNs have same structure with shared weights. We converted the sounds to 128×128 Mel spectrograms and fed them as inputs to the model.

5.3 Loudness Calibration

According to G3, the perception of loudness may vary from person to person. Thus, a user should be able to calibrate the loudness of sound sources. Ambient-Sonic supports a small configuration panel where a user can interactively achieve that (Figure 6b). The panel provides several sliders representing the volume of each sound source. The sliders are initialized with the values calculate from Section 5.2.1. The ranges for each slider is set in a way that a user can change the volume by ± 10 only. This is to ensure that users do not set a volume too loud, or too silent. Further, a user can (de)select sound sources for sonification. Initially, all the sounds are selected. The panel is available to users anytime they want although it is designed to be an optional on-boarding experience. The panel is accessible and can be controlled via a mouse or a screen reader.

5.4 Sonification

In this section, we demonstrate several use cases for Ambient-Sonic. We show how Ambient-Sonic sonifies commonly used charts. The sonifications presented below are available here: http://datasonification.herokuapp.com/examples.

Comparing scores for students. Consider the case of comparing test scores for three students. To sonify this data, Ambient-Sonic will at first select a sound source randomly from the full collection of bird sounds. It will then select

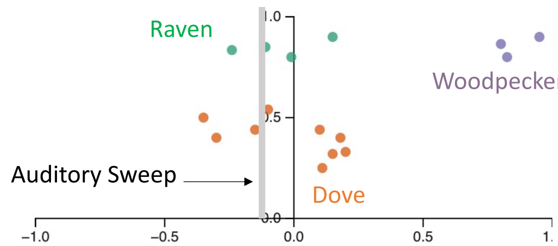


Fig. 7. Sonification of a scatter plot with three classes of points. Each class is represented with a distinct natural sound. An auditory sweep (the grey line) moves from left to right and plays the corresponding sounds whenever it intersects with a point or a group of points.

two more sounds that are most dissimilar to the first sound. Their loudness is already normalized using the peak and loudness normalization (Figure 6a).

The dataset is converted to 0-1 scale (normalized). The volume of the sound sources is set equal to the normalized score of the students. We used stereo panning to position the sound sources on left/right ear alternatively. The final sonification is a 15 sec long audio where all the sound sources are played together (G5). A user can increase or decrease the length of the sonification by pressing DOWN and UP arrow keyboard shortcuts (Figure 6d). We also provide keyboard shortcuts to interact with the sonification (Figure 6d, G6). For example, a user can press 1 to select the first student and listen to the sonification of that student only. A subsequent press on 1 will deactivate the student in the sonification. In this way, a user can listen to any combination of the datapoints (students), similar to the way a user would interact with a visualization. Pressing the ESC button will reset any selection.

Comparing trends in stock prices. Figure 2 presented how Ambient-Sonic-Lite sonified a line chart or a multi line chart. Based on Section 4, we refined this representation. First, instead of using a fixed number of segments, the number of segments is set to T/t where T is the time of the full sonification and t is the interval for changing volumes. The initial value of T and t are set to 20s and 5ms. A user can further control T and t using UP and DOWN arrow keyboard shortcuts (Figure 6d). For comparing multiple stocks together, we used the similar method described for bar chart (G5).

Locate points in 2D cartesian space. Existing tools such as Highcharts [13] and Audio Graph [1] use auditory sweeps to scan a 2D space from left to right and play the points whenever an intersect happens. However, they do not convey the location of the points in the 2D Cartesian space and relative distance between points. Moreover, they do not support sonification of points that have external group level or class, a common scenario for scatter plots. In Ambient-Sonic, a user can provide class levels for the data points. For example, Figure 7 shows a scatter plot with three classes of points (green, orange, and purple).

Ambient-Sonic uses an auditory sweep to scan the scatter plot from left to right. We use stereo panning to move the auditory sweep, a mild white noise, from left ear to right ear. Each class of points are assigned a unique natural sound. For example, in Figure 7 the three sets of points will be represented with Raven, Dove, and Woodpeckers. In the case of an intersection between the points and sweep, the corresponding natural sounds are played if they are not already playing. This ensures that the sounds do not overlap. For example, Figure 7 shows an intersection point where both Raven and Dove sound will be played together (G5). Similar to bar and line chart, the height (y value) of a point is expressed through the volume.

Table 1. Participant Information

Participant Id	Gender	Age	Musically Trained?	Highest Educational Degree	Data Analysis Experience?
P1	Male	28	Yes	Masters	Yes
P2	Female	27	No	Masters	Yes
P3	Male	31	No	Masters	Yes
P4	Female	24	No	Bachelors	No
P5	Male	22	No	Bachelors	No
P6	Female	26	No	Bachelors	Yes
P7	Male	25	Yes	Bachelors	No
P8	Male	27	No	Masters	Yes
P9	Female	25	No	Bachelors	Yes
P10	Female	33	No	Masters	No
P11	Female	28	No	High School	Yes
P12	Male	24	No	Bachelors	Yes
P13	Male	25	No	Bachelors	Yes
P14	Male	24	Yes	Bachelors	Yes

5.5 Implementation Notes

Ambient-Sonic is currently web-based. We used Flask as a Python Server and Javascript as a front end language for supporting user interactions. Howler.js [2] was used for basic audio functionalities such as play, pause, stop, changing volume, etc.

6 EVALUATION

We conducted a user study with 14 participants to evaluate Ambient-Sonic. In this study, we aimed at finding effectiveness and usability of Ambient-Sonic in comparison to existing sonification tools.

6.1 Participants

We recruited 14 participants through local mailing lists, university mailing lists, and social media posts. None of the participants participated in our pilot study. Our inclusion criteria included familiarity with rudimentary summary statistics such as *min*, *max*, and *average*, owning a stereo headphone, and having no hearing impairments; familiarity with data visualization or data analysis was not a requirement. None of the participants reported prior experience with data sonification. Participants varied in terms of gender, age, and whether or not they have prior musical training. All participants were sighted. We provide their information in Table 1. The Data Analysis column lists Yes if a participant had prior experience in inferring insights and summary statistics from a dataset and No otherwise.

6.2 Study Conditions

We conducted a repeated-measures within-subject experiment with the following two conditions.

- **C1. Baseline:** This condition included sonifications created from HighCharts [13] and Sonification Sandbox [49]. It supported interactions such as increasing or decreasing the speed of a sonification. It did not feature any interactions with the sonification as it was not supported in Highcharts and Sonification Sandbox.

[2]https://howlerjs.com

- **C2. Ambient-Sonic:** This condition included sonifications created using our tool. It supported keyboard interactions listed in Figure 6d.

We included HighCharts [3, 13] in the baseline as it is a widely used library and their sonification module provides an easy and recent API to sonify various charts. We also included Sonification Sandbox [49] in our baseline as it is open-source and provides a desktop application to create sounds. We believe together they form a fair representation of the existing musical notes based sonification. Each condition featured tasks based on Bar Charts (**V1**), and Line Charts (**V2**).

Additionally, the Ambient-Sonic condition (C2) also featured Scatter Plots (**V3**). We did not include scatter plots for C1 as Highcharts do not support finding location of points in a scatter plot. Further, there are no features to sonify multiple classes of points in Highcharts scatter plot. Similarly, Sonification Sandbox does not support scatter plots.

6.3 Tasks

We designed the following tasks for each condition. The tasks are categorized into the scenarios described in Section 5.4, corresponding to different chart types (**V1-3**). The tasks reflected the common sonification tasks discussed in section 2.3.

- **T1: Inferring test scores for students (Bar Chart).** To realize how sonification of a bar chart can be helpful, we presented the scenario of inferring test scores for *n* students from an audio. Each student essentially represents a bar in this scenario. We limit *n* to 3 to 5.
  - **T1.1.** Finding the student with minimum or maximum score. Example questions: "*Which student has the highest score?*" "*Which student has the lowest score?*"
  - **T1.2.** Comparing the scores for a pair of students. Example questions: "*Between the students represented by the raven and woodpecker sound, which student has the higher score?*" "*Between student 1 and 5, which student has the lower score?*"
- **T2: Inferring stock prices (Line Chart).** We presented the scenario of understanding stock prices to evaluate sonificaion of line charts. The charts featured 1 or 2 lines.
  - **T2.1.** Understanding the trend of a single stock. Example questions: "*What is the overall trend for the stock price?*" "*Which of the following is true about the stock price?*" We provided multiple choices for these questions. See supplemental materials for a detail list of questions and multiple choices.
  - **T2.2.** Forecast trend of a single stock. Example question: "*Based on the audio, what will be your forecast for the stock price in near future?*" See supplemental materials for the multiple choices.
  - **T2.3.** Compare trends for two stocks. Example questions: "*Did the price of any of the stocks face a major drop?*" "*Which of the following is true about the prices of the stocks in recent years?*"

For line charts with two stocks, we intentionally avoided asking questions that require answers in terms of referring to an individual stock. This is because Highcharts do not assign different sound sources to different lines; the same sound is played for both stocks with pitch proportionate to the lines. Thus, there is no meaningful way to refer to the stocks. Finally, we designed the following tasks for scatter plots, although they did not feature in the baseline.

- **T3: Identify location of points (Scatter Plot).**
  - **T3.1.** Find stereo panning for points. Example question: "*Can you tell which side of your ear the crickets are buzzing?*"
  - **T3.2.** Identify relative distance between points. "*Which pair of points have the shorter average distance between them?*"

The underlying datasets for each condition featured similar complexity. Example sounds used for both baseline and Ambient-Sonic are available here: http://datasonification.herokuapp.com/examples.

6.4 Protocol

Similar to our pilot study, we deployed Ambient-Sonic on a public server and conducted the sessions in ZOOM. Each session started with participants signing the consent form. We then shared the URL for the study. To minimize the learning effect, we counterbalanced the ordering of study conditions and visualization. Scatter plot (V3) was not included in the counterbalancing process as it was only available for C2. V3 was always put at the end of C2. We maintained the same task order for each visualization as they are presented with increasing complexity to the participants.

The condition C1 was divided into two parts, containing sounds and questions for V1 and V2. For each chart type, we at first provided participants with 2 training sounds. We provided verbal instruction on how to interpret the sounds. Once participants felt comfortable with the training sonifications, we provided 3-4 test sounds, one after one. For each test sound, we asked questions relevant to the tasks listed above. We followed the same protocol for C2, except it featured an on-boarding interface similar to 6b where a participant could listen to individual sound sources and calibrate their volumes using the method described in Section 5.3.

Each session lasted for 1.5 hours. The experimenter took notes during the session. All sessions were video recorded and transcribed. Each session culminated with participants making suggestions and recommendations. At the end of a session, participants received a 15 dollars Amazon gift card.

6.5 Results

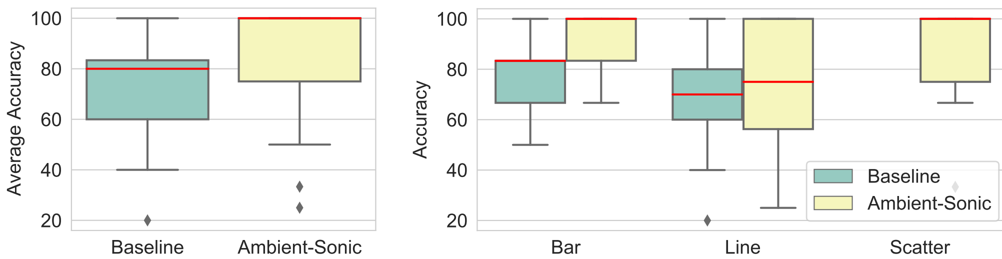
To measure performance, we calculated participants' accuracy in answering the questions. Participants also rated the study conditions on a Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree) based on five subjective measures. We performed paired t-test (two tailed) with Bonferroni Correction to measure statistical significance.

6.5.1 Accuracy/Performance. We found that Ambient-Sonic has a statistically significant impact on assisting participants answer the questions. Figure 8a presents participants' average accuracy in answering the questions. While using Ambient-Sonic, participants were able to answer the questions with an average accuracy of 85.21% ($Median = 84.62, Mode = 92.3, SD = 12.14$). In contrast, while using the baseline the average accuracy was 72.72% ($Median = 72.72, Mode = 63.63, SD = 14.47$). The difference (12.49) was statistically significant ($t = 2.76, df = 13, p < 0.02$).

Figure 8b further presents participants' average accuracy in answering the questions for bar, line, and scatter plots. For bar chart, the average accuracy were 92.32% ($Median = 100, , Mode = 100, SD = 6.79$) with Ambient-Sonic and 78.21% ($Median = 83.33, Mode = 83.33, SD = 13.63$) with baseline. The difference (14.11) was statistically significant ($t = 2.51, df = 13, p < 0.03$). For line chart, the average accuracy were 75% ($Median = 75, Mode = 72, SD = 32.33$) with Ambient-Sonic and 66.42% ($Median = 60, Mode = 60, SD : 25.61$) with baseline. The difference (8.58) was not statistically significant.

Finally, while using Ambient-Sonic participants answered the questions for scatter plots with an average accuracy of 87.18% ($Median = 100, Mode = 100, SD = 27.88$).

6.5.2 Subjective Measures and Qualitative Feedback. Figure 8c presents participants ratings on five subjective measures. Participants rated Ambient-Sonic favorably for the measure *Distinguishability.* This measure asked participants to rate the conditions based on how well they can identify multiple categories from the audio. Paricipants understood the



(a) (b)

Subjective Measure	Baseline (Participant Rating)					Ambient-Sonic(Participant Rating)				
	1	2	3	4	5	1	2	3	4	5
Distinguishability	0	4	3	5	1	0	1	3	4	6
Enjoyment	0	3	3	3	5	0	3	3	4	4
Mental Demand	2	2	7	2	1	3	6	3	2	0
Easiness	0	2	5	6	1	0	3	2	6	3
Annoyance/Frustration	3	5	3	2	0	4	6	2	1	0

(c)

Fig. 8. **Study Results.** (a) Participants' average accuracy in answering inference questions while using Ambient-Sonic and the baseline. Red lines show the median values. (b) Participants' average accuracy in answering inference questions for bar, line, and scatter plots. (c) Subjective Measures.

presence of multiple categories of a nominal data easily through Ambient-Sonic. In contrast, participants struggled to distinguish multiple stocks when using Highcharts. P7 said

> *It was easy when there was one stock (using Highcharts). But it was difficult to track two stocks. I get that there are two frequencies all the time. At one point, I think one of the stocks decreased. But, which one? I could not tell which one saw the drop in price. I started over the sound to assign numbers to the frequencies and then track them. Even then it was difficult.*

P5 mentioned that Ambient-Sonic made *it easier to track the two stocks.* P1 mentioned that they could easily distinguish the students using Ambient-Sonic since *it was easy to identify the different bird sounds and track their changes.* Interactions also helped participants separate sound sources and compare them pair wise (P5, P8). It is worth noting here that although the performance of participants did not improve significantly while using line charts in Ambient-Sonic, they employed less efforts while answering the questions for line charts in Ambient-Sonic.

Participants rated both conditions equally on the metric *Enjoyment.* They expressed that both musical notes and natural sounds are *fun, and interesting* (P2-3, P8). Additionally, natural sounds evoked several emotions from the participants. P1 said it has *the essence of sitting outside in a park.* P9 said *I thought I was listening to birds from my balcony.*

Participants enjoyed spatial sounds in Ambient-Sonic scatter plots. P10 referred to the scatter plot sound as *cool and innovative*.

Finally, participants rated Ambient-Sonic slightly better in terms of how *frustrated* and *annoyed* they felt while answering the questions. Overall, the general consensus was that Ambient-Sonic is easy to use and did not warrant excessive cognitive work.

7 DISCUSSION AND FUTURE WORK

Impact on sonification research. Sonification research has diverse applications— from understanding non-traditional data such as protein structure to helping a blind person navigate daily life, and access data visualization (discussed in Section 2). Our visual system can consume a large number of data quickly whereas our auditory senses are usually limited to perceive a small amount of data. Thus, we need methods to overcome, or adapt to the limitation of our auditory senses. In this paper, we made a small stride towards that goal. In particular, we show for the first time that natural sounds can be used to represent data. Further, they can be blended together to increase information flux in a sonification. Although our method is not tied to any specific application area at this moment, we believe it will work as a catalyst for further research on using natural sounds for sonification.

We are not so foolish to believe that our method will replace the existing musical note-based sonification tools. Rather we believe natural sounds can augment the existing tools. The existing tools can be furnished with the support of natural sounds and then based on user preference these tools can either use musical notes or natural sounds for sonification. Our future work will concentrate on how the lessons learned from this work can be transferred to different application areas.

Improvements. We used volume to quantify quantitative values, in contrast to popularly used pitch. While participants performance shows that volume was effective in conveying data, future works may concentrate on evaluating different auditory properties discussed in section 2.4 for this purpose.

The sound sources used in this paper are collected from freely available web contents. However, these sounds are not created or collected to be used for data sonification. We believe a dataset of natural sound sources where the sounds are collected for data sonification will produce a better mix and increase pleasantness of the overall sound.

Synthesizing natural sounds into a single mix was a major challenge in this work. We outlined several key design guidelines for mapping data to natural sounds. We also introduced several automated methods for loudness normalization, sound similarity, etc. To increase the quality of blending, we experimented with several neural networks. For example, we experimented with Variational Autoencoders to learn the underlying latent representation of natural sounds. Our goal was to learn the latent structure so that we could produce these sounds without any external noise. However, the results were not convincing. Future work may explore a more advanced model such as a hierarchical recurrent variational autoencoder for learning latent spaces [35].

While the parallel sonification mechanism of Ambient-Sonic was effective in conveying nominal data, we limit the number of categories to at most 5 since our mind can only concentrate on a limited number of unique sounds together (i.e., the bandwidth problem). Future work may concentrate on finding ways to scale this mechanism to beyond 5 categories.

Future Work. Besides sighted people, we believe Ambient-Sonic can benefit different target populations. For instance, we plan to conduct a study with people with vision impairments to understand the feasibility and challenges of Ambient-Sonic as a standalone sonification technique for this population. We also believe that people who lack graph literacy, i.e., the ability to understand information that is presented graphically [39], can benefit from Ambient-Sonic. Therefore, it is worthy of conducting a study with this target population to understand how they can extract information and make inferences with Ambient-Sonic.

8 CONCLUSION

In this paper, we presented a new sonification technique using natural sounds. Our work is grounded in the observation that natural sounds are integrated in our day-to-day life, are easily distinguishable, and have hedonic values for meditation, and well-being. We validated these observations in a pilot study and found that natural sounds have the potential to turn serial cosumption of data to parallelly listening to multiple categories of a nominal data. Informed by the study, we designed Ambient-Sonic that uses several intelligent audio processing functionalities to blend multiple natural sounds to a coherent single sound. To evaluate the tool, we conducted a summative study. The outcome of the study suggests Ambient-Sonic improved participant performance in inferring summary statistics and trends from a dataset.

We believe this work suggests many interesting future avenues of research. For one thing, Sonification is a multi-disciplinary research area and we believe our work provides important ground works for further research in HCI, Accessibility, and ICAD community. We believe our work will motivate future efforts to use natural sounds for sonification.

REFERENCES

[1] 2021. Audio Graphs. Available at https://developer.apple.com/documentation/accessibility/audio_graphs.
[2] 2021. Bring accessibility to charts in your app. Available at https://developer.apple.com/videos/play/wwdc2021/10122/.
[3] 2021. Highcharts: Interactive Javascript Charts Library. Available at https://www.highcharts.com.
[4] 2021. Twotone: an open source data sonification web app. Available at https://github.com/datavized/twotone.
[5] Jesper J Alvarsson, Stefan Wiens, and Mats E Nilsson. 2010. Stress recovery during exposure to nature sound and environmental noise. *International journal of environmental research and public health* 7, 3 (2010), 1036–1046.
[6] Mark Ballora, Bruce Pennycook, Plamen C Ivanov, Leon Glass, and Ary L Goldberger. 2004. Heart rate sonification: A new approach to medical diagnosis. *Leonardo* 37, 1 (2004), 41–46.
[7] Stephen Barrass. 1998. *Auditory information design.* The Australian National University (Australia).
[8] Stephen Barrass. 2005. A perceptual framework for the auditory display of scientific data. *ACM Transactions on Applied Perception (TAP)* 2, 4 (2005), 389–402.
[9] Andrea Batch, Biswaksen Patnaik, Moses Akazue, and Niklas Elmqvist. 2020. Scents and Sensibility: Evaluating Information Olfactation. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.
[10] Marc G Berman, John Jonides, and Stephen Kaplan. 2008. The cognitive benefits of interacting with nature. *Psychological science* 19, 12 (2008), 1207–1212.
[11] Lorna M Brown, Stephen A Brewster, SA Ramloll, R Burton, and Beate Riedel. 2003. Design guidelines for audio presentation of graphs and tables. International Conference on Auditory Display.
[12] Rachel T. Buxton, Amber L. Pearson, Claudia Allou, Kurt Fristrup, and George Wittemyer. 2021. A synthesis of health benefits of natural sounds and their distribution in national parks. *Proceedings of the National Academy of Sciences* 118, 14 (2021). https://doi.org/10.1073/pnas.2013097118 arXiv:https://www.pnas.org/content/118/14/e2013097118.full.pdf
[13] Stanley J Cantrell, Bruce N Walker, and Øystein Moseng. 2021. Highcharts Sonification Studio: An Online, Open-Source, Extensible, And Accessible Data Sonification Tool. *The 26th International Conference on Auditory Display* (2021). Virtual Conference (25-28 June).
[14] Benjamin K. Davison and B. Walker. 2007. Sonification Sandbox Reconstruction: Software Standard for Auditory Graphs.
[15] Wanda L Diaz Merced. 2013. *Sound for the exploration of space physics data.* Ph.D. Dissertation. University of Glasgow.
[16] Stephanie Elzer, Edward Schwartz, Sandra Carberry, Daniel Chester, Seniz Demir, and Peng Wu. 2008. Accessible bar charts for visually impaired users. In *Fourth Annual IASTED Intl. Conf. on Telehealth and Assistive Technologies*. 55–60.
[17] Leo Ferres, Gitte Lindgaard, Livia Sumegi, and Bruce Tsuji. 2013. Evaluating a tool for improving accessibility to charts and graphs. *ACM Transactions on Computer-Human Interaction (TOCHI)* 20, 5 (2013), 1–32.
[18] Sean Follmer, Daniel Leithinger, Alex Olwal, Akimitsu Hogge, and Hiroshi Ishii. 2013. inFORM: dynamic physical affordances and constraints through shape and object actuation.. In *Uist*, Vol. 13. 2501988–2502032.

[19] Susan N Friel, Frances R Curcio, and George W Bright. 2001. Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in mathematics Education* 32, 2 (2001), 124–158.

[20] Cagatay Goncu, Anuradha Madugalla, Simone Marinai, and Kim Marriott. 2015. Accessible on-line floor plans. In *Proceedings of the 24th international conference on World Wide Web*. 388–398.

[21] Thomas Hermann, Andy Hunt, and John G Neuhoff. 2011. *The sonification handbook*. Logos Verlag Berlin.

[22] Leona Holloway, Kim Marriott, and Matthew Butler. 2018. Accessible maps for the blind: Comparing 3D printed models with tactile graphics. In *Proceedings of the 2018 chi conference on human factors in computing systems*. 1–13.

[23] Shaun K Kane and Jeffrey P Bigham. 2014. Tracking stemxcomet: teaching programming to blind students via 3D printing, crisis management, and twitter. In *Proceedings of the 45th ACM technical symposium on Computer science education*. 247–252.

[24] Rachel Kaplan. 2001. The nature of the view from home: Psychological benefits. *Environment and behavior* 33, 4 (2001), 507–542.

[25] Stephen Kaplan. 1995. The restorative benefits of nature: Toward an integrative framework. *Journal of environmental psychology* 15, 3 (1995), 169–182.

[26] Fredrik Kilander and Peter Lönnqvist. 2002. A whisper in the woods-an ambient soundscape for peripheral awareness of remote processes. Georgia Institute of Technology.

[27] Gregory Koch, Richard Zemel, and Ruslan Salakhutdinov. 2015. Siamese neural networks for one-shot image recognition. *Proceedings of the 32nd International Conference on Machine Learning*.

[28] Mathieu Le Goc, Charles Perin, Sean Follmer, Jean-Daniel Fekete, and Pierre Dragicevic. 2018. Dynamic composite data physicalization using wheeled micro-robots. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 737–747.

[29] Bradley S Mauney and Bruce N Walker. 2004. Creating functional and livable soundscapes for peripheral monitoring of dynamic data. Georgia Institute of Technology.

[30] Joachim Meyer. 2000. Performance with tables and graphs: Effects of training and a visual search model. *Ergonomics* 43, 11 (2000), 1840–1865.

[31] NASA. [n.d.]. Data Sonification: A New Cosmic Triad of Sound. https://www.nasa.gov/mission_pages/chandra/news/data-sonification-a-new-cosmic-triad-of-sound.html

[32] Biswaksen Patnaik, Andrea Batch, and Niklas Elmqvist. 2018. Information olfactation: Harnessing scent to convey data. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 726–736.

[33] Graphical Perception. 1984. Theory, Experimentation, and Application to the Development of Graphical Methods William S. Cleveland; Robert McGill. *J. Amer. Statist. Assoc.* 79, 387 (1984), 531–554.

[34] Eleanor Ratcliffe, Birgitta Gatersleben, and Paul T Sowden. 2013. Bird sounds and their contributions to perceived attention restoration and stress recovery. *Journal of Environmental Psychology* 36 (2013), 221–228.

[35] Adam Roberts, Jesse Engel, Colin Raffel, Curtis Hawthorne, and Douglas Eck. 2018. A hierarchical latent vector model for learning long-term structure in music. In *International conference on machine learning*. PMLR, 4364–4373.

[36] Kavous Salehzadeh Niksirat, Chaklam Silpasuwanchai, Mahmoud Mohamed Hussien Ahmed, Peng Cheng, and Xiangshi Ren. 2017. A framework for interactive mindfulness meditation using attention-regulation process. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2672–2684.

[37] Arvind Satyanarayan, Dominik Moritz, Kanit Wongsuphasawat, and Jeffrey Heer. 2017. Vega-lite: A grammar of interactive graphics. *IEEE transactions on visualization and computer graphics* 23, 1 (2017), 341–350.

[38] Nik Sawe, Chris Chafe, and Jeffrey Treviño. 2020. Using Data Sonification to Overcome Science Literacy, Numeracy, and Visualization Barriers in Science Communication. *Frontiers in Communication* 5 (2020), 46. https://doi.org/10.3389/fcomm.2020.00046

[39] Priti Shah and Eric G Freedman. 2011. Bar and line graph comprehension: An interaction of top-down and bottom-up processes. *Topics in cognitive science* 3, 3 (2011), 560–578.

[40] Daniel R Smith and Bruce N Walker. 2005. Effects of auditory context cues and training on performance of a point estimation sonification task. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition* 19, 8 (2005), 1065–1087.

[41] Tony Stockman. 2010. Listening to people, objects and interactions. In *Human Interaction with Auditory Displays–Proceedings of the Interactive Sonification Workshop*. 3–10.

[42] Faisal Taher, John Hardy, Abhijit Karnik, Christian Weichel, Yvonne Jansen, Kasper Hornbæk, and Jason Alexander. 2015. Exploring interactions with physically dynamic bar charts. In *Proceedings of the 33rd annual acm conference on human factors in computing systems*. 3237–3246.

[43] Faisal Taher, Yvonne Jansen, Jonathan Woodruff, John Hardy, Kasper Hornbæk, and Jason Alexander. 2016. Investigating the use of a dynamic physical bar chart for data exploration and presentation. *IEEE transactions on visualization and computer graphics* 23, 1 (2016), 451–460.

[44] Petra Vetter, Łukasz Bola, Lior Reich, Matthew Bennett, Lars Muckli, and Amir Amedi. 2020. Decoding natural sounds in early "visual" cortex of congenitally blind individuals. *Current Biology* 30, 15 (2020), 3039–3044.

[45] Karen Vines, Chris Hughes, Laura Alexander, Carol Calvert, Chetz Colwell, Hilary Holmes, Claire Kotecki, Kaela Parks, and Victoria Pearson. 2019. Sonification of numerical data for education. *Open Learning: The Journal of Open, Distance and e-Learning* 34, 1 (2019), 19–39. https://doi.org/10.1080/02680513.2018.1553707

[46] Bruce Walker and Joshua Cothran. 2003. Sonification Sandbox: A graphical toolkit for auditory graphs.

[47] Bruce N Walker. 2002. Magnitude estimation of conceptual data dimensions for use in sonification. *Journal of experimental psychology: Applied* 8, 4 (2002), 211.

[48] Bruce N Walker. 2007. Consistency of magnitude estimations with conceptual data dimensions used for sonification. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition* 21, 5 (2007), 579–599.

[49] Bruce N Walker and Joshua T Cothran. 2003. Sonification Sandbox: A graphical toolkit for auditory graphs. Georgia Institute of Technology.

[50] Cheng Xu, Ali Israr, Ivan Poupyrev, Olivier Bau, and Chris Harrison. 2011. Tactile display for the visually impaired using TeslaTouch. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*. 317–322.

[51] Chi-Hua Yu and Markus J Buehler. 2020. Sonification based de novo protein design using artificial intelligence, structure prediction, and analysis using molecular modeling. *APL bioengineering* 4, 1 (2020), 016108.

[52] Chi-Hua Yu, Zhao Qin, Francisco J Martin-Martinez, and Markus J Buehler. 2019. A self-consistent sonification method to translate amino acid sequences into musical compositions and application in protein design using artificial intelligence. *ACS nano* 13, 7 (2019), 7471–7482.

[53] Haixia Zhao, Catherine Plaisant, Ben Shneiderman, and Ramani Duraiswami. 2004. Sonification of Geo-Referenced Data for Auditory Information Seeking: Design Principle and Pilot Study.. In *ICAD*.

[54] Haixia Zhao, Catherine Plaisant, Ben Shneiderman, and Jonathan Lazar. 2008. Data sonification for users with visual impairment: a case study with georeferenced data. *ACM Transactions on Computer-Human Interaction (TOCHI)* 15, 1 (2008), 1–28.