“What Makes Sonification User-Friendly?” Exploring Usability and User-Friendliness of Sonified Responses

Ather Sharif*
asharif@cs.washington.edu
Paul G. Allen School of Computer Science & Engineering | DUB Group, University of Washington
Seattle, Washington, USA

Olivia H. Wang*
wang4@cs.washington.edu
Paul G. Allen School of Computer Science & Engineering,
University of Washington
Seattle, Washington, USA

Alida T. Muongchan
alidatm@uw.edu
Human Centered Design and Engineering,
University of Washington
Seattle, Washington, USA

ABSTRACT

Sonification is a commonly used technique to make online data visualizations accessible to screen-reader users through auditory means. While current sonification solutions provide plausible utility (usefulness) to screen-reader users in exploring data visualizations, they are limited in exploring the quality (usability) of the sonified responses. In this preliminary exploration, we investigated the usability and user-friendliness of data visualization sonification for screen-reader users. Specifically, we evaluated the *Pleasantness*, *Clarity*, *Confidence*, and *Overall Score* of discrete and continuous sonified responses generated using various oscillator waveforms and synthesizers through user studies with 10 screen-reader users. Additionally, we examined these factors using both simple and complex trends. Our results show that screen-reader users preferred distinct non-continuous responses generated using oscillators with *square* waveforms. We utilized our findings to extend the functionality of `Sonifier`—an open-source JavaScript library that enables developers to sonify online data visualizations. Our follow-up interviews with screen-reader users identified the need to personalize the sonified responses per their individualized preferences.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in visualization**; **Accessibility systems and tools**; **Empirical studies in accessibility**.

KEYWORDS

sonification, audio graphs, waveforms, visualizations, screen-reader users

ACM Reference Format:
Ather Sharif, Olivia H. Wang, and Alida T. Muongchan. 2022. “What Makes Sonification User-Friendly?” Exploring Usability and User-Friendliness of Sonified Responses. In *The 24th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS ’22), October 23–26, 2022, Athens, Greece.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3517428.3550360

*These authors contributed equally to this work.

ASSETS ’22, October 23–26, 2022, Athens, Greece

ACM ISBN 978-1-4503-9258-7/22/10.
https://doi.org/10.1145/3517428.3550360

1 INTRODUCTION

With the increasing data representation through visualizations comes the critical need to make these visualizations accessible to people who may not be able to extract information using visual means (e.g., screen-reader users) [6, 16, 18, 21]. According to recent findings from Sharif *et al.*, due to the inaccessibility of visualizations, screen-reader users extract information 62% less accurately and spend 211% more interaction time with online data visualizations compared to non-screen-reader users [21]. Researchers and developers have implemented several approaches and strategies to make online data visualizations accessible [8, 10, 15, 22, 24, 25]. Among these techniques is sonification, often referred to as “audio graphs,” a widely-used approach in conveying data through auditory channels to screen-reader users.

Several prior works have utilized sonification to improve the accessibility of online data visualizations [2, 3, 7, 13, 14, 19, 24, 26, 29]. However, current solutions are focused on the utility (usefulness) of sonification to screen-reader users and provide limited insights into the quality (usability and user-friendliness) of the sonified responses. Therefore, in this work, we sought to examine and improve the usability and user-friendliness of sonified responses generated from online data visualizations created using JavaScript libraries (e.g., D3 [4]). Prior research has explored the “pleasantness” of sonified responses [1, 9, 20]. However, the most relevant research to our work is the recent exploration by Wang *et al.* [27], in which they examined the impact of various auditory channels (e.g., pitch, volume) on users’ perception of data and visualization. We build on their work by investigating the effects of different oscillator waveforms and synthesizers on the pleasantness and users’ confidence in interpreting simple and complex sonified responses.

To perform our investigation, we developed several sonification prototypes using the `Tone.js` library [17], incorporating different configurations for *Sound Types* (oscillator waveforms and synthesizers), *Continuity Levels* (interval for sounds between data points), and *Trend Types* (simple and complex). We finalized six prototypes through Wizard-of-Oz [5, 12] and pilot studies with users, employing a user-centered iterative design process. We evaluated these prototypes by collecting subjective ratings for *Pleasantness*, *Clarity*, *Confidence*, and *Overall Score* through user studies with 10 screen-reader users. We utilized our findings to extend the functionalities of `Sonifier`—a recently-introduced open-source JavaScript library that generates sonified responses for two-dimensional single-series data [23, 24]. Additionally, we conducted follow-up interviews with our participants to solicit feedback on improving their experiences with sonified responses.

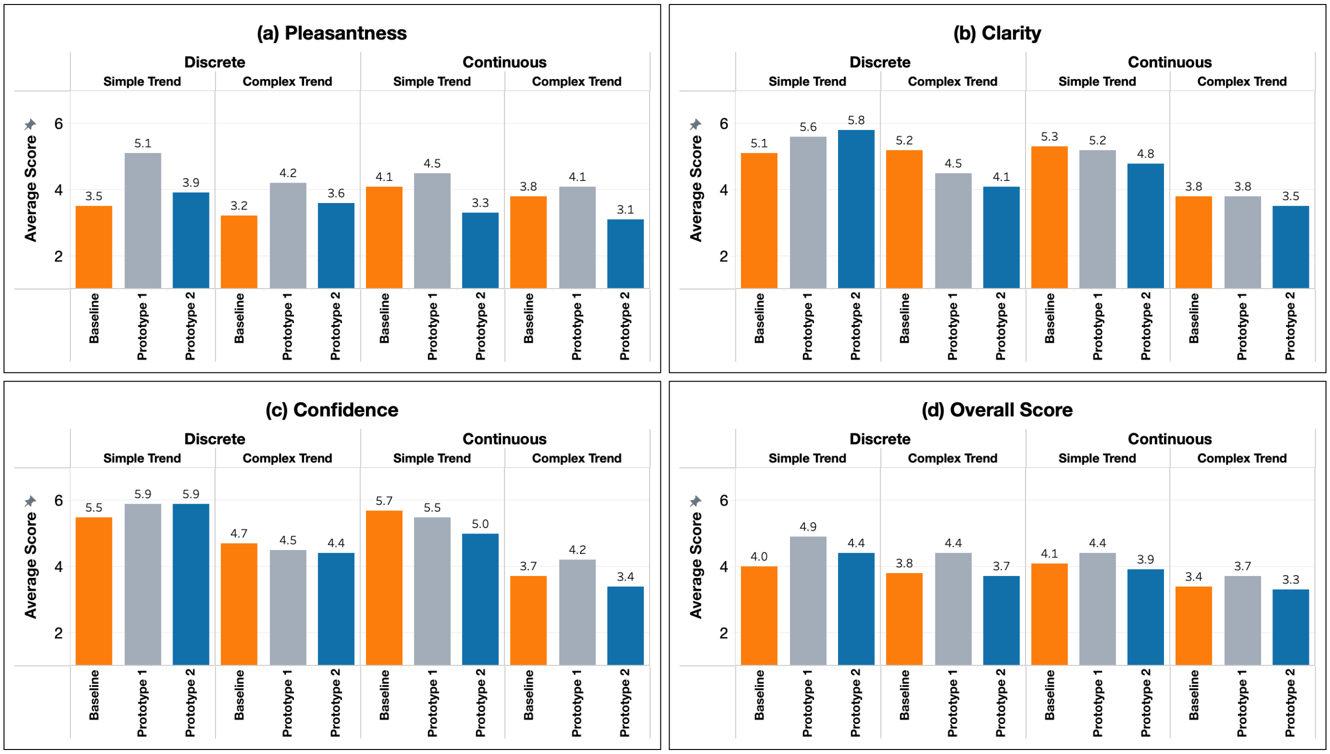

Figure 1: Average scores per Continuity Level and Trend Type for each Sound Type for (a) *Pleasantness*, (b) *Clarity*, (c) *Confidence*, and (d) *Overall Scores*. *Baseline* was a *sawtooth* waveform OmniOscillator, *Prototype 1* was a *square* waveform OmniOscillator, and *Prototype 2* was a *square* waveform MonoSynthesizer. Subjective ratings were collected using a Likert scale ranging from 1 (worst) to 7 (best).

In this work, we contribute the empirical findings from our preliminary explorations. Specifically, we provide the quantitative and qualitative results from our user studies with 10 screen-reader users. Additionally, we provide enhancements to the open-source *Sonifier* library and identify avenues for future work.

2 USER STUDY

To assess the usability and user-friendliness of sonified responses, we conducted a user study with 10 screen-reader users, subsequently interviewing them to gather insights on further improving the sonified responses.

2.1 Prototypes

We created several sonification prototypes employing different combinations for oscillator waveforms, synthesizers, and partial counts (number of harmonics to generate the waveform). We developed these prototypes using the `Tone.js` JavaScript library [17]—a widely-used framework to generate sounds in the browser. Then, we eliminated the prototypes that users found undesirable through Wizard-of-Oz [5, 12] and pilot studies. After elimination, our sound types comprised an oscillator (termed "OmniOscillator" by `Tone.js`) and a synthesizer (termed "MonoSynthesizer" by `Tone.js`), with both using the *square* waveform and a partial count of *8*. We created a continuous (minimal time interval between sounds for each data point) and discrete (clear time interval between sounds for each data point) prototype for each sound type. We used the default settings from the `Sonifier` library as our baseline measure. Since the baseline only supports discrete responses, we developed our fifth prototype that generated a continuous response for the baseline to account for balanced conditions. Our final set contained 12 prototypes (including the baseline), six for each *Trend Type* (simple and complex).

2.2 Participants & Procedure

Our participants were 10 screen-reader users (Appendix A, Table 1). Four identified as women, five as men, and one as non-binary. Their average age was 48.4 (*SD*=14.4) years. We compensated our participants with a $10 Amazon gift card for 30 minutes of their time. We supervised our user studies online using Zoom.

For each participant, we played six sonified responses (five prototypes + baseline) for simple trends, randomizing the order to account for learning effect. At the end of each sonification, we collected their subjective ratings for *Pleasantness* (timbre), *Clarity* (assessment of the sound to identify the trend clearly), *Confidence* (user's confidence in understanding the overall trend), and *Overall Score* (subjective assessment of the sound overall, including any factors not mentioned above). We used a Likert scale for subjective ratings, ranging from 1 (worst) to 7 (best). Then, we asked follow-up questions from our users to gather insights on the areas of improvement. We followed the same steps for complex trends. Our study sessions, on average, took approximately 30 minutes from start to finish.

2.3 Design & Analysis

The experiment was a 3×2×2 within-subjects design with the following factors and levels:

- *Sound Type (S)*: {Baseline, OmniOscillator, MonoSynthesizer}
- *Continuity Level (C)*: {discrete, continuous}
- *Trend Type (T)*: {simple, complex}

We used *Pleasantness* (PL), *Clarity* (CL), *Confidence* (CF), and *Overall Score* (OS) as our dependent variables. In our analysis, all our dependent variables were ordinal (on a scale of 1-7). Therefore, we conducted our analysis using an ordinal logistic regression [11, 28]. We also included $Subject_r$ as a random factor to account for repeated measures. We tested our participants over 3×2×2=12 conditions, resulting in 12×10=120 total trials.

2.4 Results

We present the results of our user studies assessing *Pleasantness* (PL), *Clarity* (CL), *Confidence* (CF), and *Overall Score* (OS) of sonified responses. Additionally, we present our findings on areas of improvement for sonified responses from our follow-up interviews.

2.4.1 Pleasantness (PL). Our results show a significant main effect of S ($\chi^2(2, N{=}10){=}12.15$, $p{<}.05$, Cramer's $V{=}.42$) on *PL* overall, indicating that *PL* was significantly different between the three *S* groups. Overall, the sonified responses from *OmniOscillator* with *square* waveform and discrete continuity level outperformed the other prototypes (M=4.65). C ($p \approx .844$) and T ($p \approx .122$) did not have a significant main effect on *PL*. Figure 1 and Table 2 show average *PL* scores for each independent variable.

2.4.2 Clarity (CL). We found a significant main effect of C ($\chi^2(1, N{=}10){=}11.03$, $p{<}.001$, Cramer's $V{=}.40$) and T ($\chi^2(1, N{=}10){=}22.83$, $p{<}.001$, Cramer's $V{=}.57$) on *CL* overall. This result indicates that *CL* was significantly different between discrete and continuous sounds and also between simple and complex trends. On average, the *Baseline* had the best scores (M=5.15). S ($p \approx .476$) did not have a significant main effect on *CL*. We show the average scores for *CL* in Figure 1 and Table 2 for each independent variable.

2.4.3 Confidence (CF). C ($\chi^2(1, N{=}10){=}6.36$, $p{<}.05$, Cramer's $V{=}.30$) and T ($\chi^2(1, N{=}10){=}42.34$, $p{<}.001$, Cramer's $V{=}.78$) had a significant main effect on *CF*, indicating that *CF* significantly differed between simple and complex trends as well as discrete and continuous sounds. Similar to *PL*, the sonified responses from *OmniOscillator* with *square* waveform and discrete continuity level had the best overall average scores (M=5.20). We did not find a significant effect of S on *CF* ($p \approx .726$). The average scores for *CF* are shown in Figure 1 and Table 2 for each independent variable.

2.4.4 Overall Score (OS). Our results show a significant main effect of T on *OS* overall ($\chi^2(1, N{=}10){=}7.16$, $p{<}.05$, Cramer's $V{=}.32$). This finding indicates that *OS* significantly varied between the simple and complex trends. *OmniOscillator* with *square* waveform and discrete continuity level performed the best compared to the other prototypes (M=4.65) on average, similar to *PL* and *CF*. We only found a marginal effect of S ($p \approx .090$) and C ($p \approx .098$) on *OS*. We display the average scores for *OS* in Figure 1 and Table 2 for each independent variable.

2.5 Areas of Improvement

Our participants recognized three areas of improvement for sonification: (1) *Personalization* (customizing the auditory output for factors including speed and frequency); (2) *Identification of extrema/outliers* (identifying the maximum and minimum data points); and (3) *Multimodality* (using a combination of different instrument sounds together with different frequencies to amplify the distinctions between data points). Future work can incorporate our findings to improve the usability and user-friendliness of sonified responses for screen-reader users.

3 `SONIFIER` LIBRARY ENHANCEMENTS

`Sonifier` is an open-source JavaScript library that generates a sonified response from two-dimensional single-series data, recently developed by Sharif *et al.* [23, 24]. Utilizing the findings from our user studies with screen-reader users, we enhanced the `Sonifier` library by (1) improving its scalability to support several sound types; and (2) modifying its default settings to those from our best-performing prototype.

We refactored and modularized the code to enable developers to create sonified responses using additional sound types, including *MonoSynthesizers* and *Envelopes*. (Currently, the `Sonifier` library only supports the creation of sonification using an OmniOscillator.) We further improve the customization for OmniOscillator by adding more configuration options, including "sourceType" (source of the oscillator; e.g., *am*), "baseType" (waveform of the oscillator; e.g., *sawtooth*), and "partialCount" (number of harmonics to generate the waveform, ranging from 1-32).

Additionally, we modified the default settings of the library based on our findings. Specifically, we changed "sourceType" to *am*, "baseType" to *square*, and "partialCount" to *8*. We made our code publicly available at the `Sonifier` library's open-source repository [23].

4 DISCUSSION & CONCLUSION

In this preliminary exploration, we investigated the usability and user-friendliness of sonified responses for screen-reader users via user studies. Our results show that screen-reader users preferred distinct non-continuous sonified responses generated using oscillators with *square* waveforms. Additionally, our follow-up interviews showed that our participants identified the need to personalize the sonified responses. We utilized our findings to enhance the capabilities of `Sonifier` [23]—an open-source library that generates sonified responses from two-dimensional data.

While the prototype generated using an oscillator with *square* waveform overall outperformed the rest of the prototypes, interestingly, the *Baseline* (using a *sawtooth* waveform) had the best average scores for *Clarity*. (However, *Clarity* did not have a statistically significant on *Sound Type*.) Our results identify avenues for future work to investigate the effects of different waveforms on different aspects of user experiences. Additionally, our findings revealed that our participants preferred to personalize the sonified responses, depending on factors such as trend type and data cardinality. Future work can utilize our findings to generate a central configuration system, enabling screen-reader users to customize the responses per their needs.

Our work identifies avenues for future research to improve the quality of sonified responses. Specifically, our work highlights that utilizing the knowledge from domains outside of Human-Computer Interaction, such as Music Theory, can improve the user experiences for people who rely on auditory channels to extract information from online data visualizations in a personalized manner. We hope our work will inspire researchers to conduct an in-depth investigation of the usability and user-friendliness of sonified responses and improve the experiences of screen-reader users with the sonification of online data visualizations.

ACKNOWLEDGMENTS

This work was supported by the University of Washington Center for Research and Education on Accessible Technology and Experiences (CREATE). We thank Jacob O. Wobbrock and Katharina Reinecke for their helpful comments and suggestions. Finally, we thank and remember our recently-departed team member Zoey for her feline support, without which the *purr*usal of this work would not have been as effective. May she cross the rainbow bridge in peace and find her way to cat heaven.

REFERENCES

[1] Dragan Ahmetovic, Federico Avanzini, Adriano Baratè, Cristian Bernareggi, Gabriele Galimberti, Luca A Ludovico, Sergio Mascetti, and Giorgio Presti. 2019. Sonification of rotation instructions to support navigation of people with visual impairment. In *2019 IEEE International Conference on Pervasive Computing and Communications (PerCom*. IEEE, 1–10.

[2] Dragan Ahmetovic, Cristian Bernareggi, João Guerreiro, Sergio Mascetti, and Anna Capietto. 2019. Audiofunctions. web: Multimodal exploration of mathematical function graphs. In *Proceedings of the 16th International Web for All Conference*. 1–10.

[3] Apple. n.d.. Audio Graphs | Apple Developer Documentation. https://developer.apple.com/documentation/accessibility/audio_graphs. (Accessed on 08/01/2021).

[4] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. 2011. D^3 data-driven documents. *IEEE transactions on visualization and computer graphics* 17, 12 (2011), 2301–2309.

[5] Nils Dahlbäck, Arne Jönsson, and Lars Ahrenberg. 1993. Wizard of Oz studies—why and how. *Knowledge-based systems* 6, 4 (1993), 258–266.

[6] Joel J Davis. 2002. Disenfranchising the Disabled: The Inaccessibility of Internet-Based Health Information. *Journal of Health Communication* 7, 4 (2002), 355–367. https://doi.org/10.1080/10810730290001701

[7] Danyang Fan, Alexa Fay Siu, Wing-Sum Adrienne Law, Raymond Ruihong Zhen, Sile O'Modhrain, and Sean Follmer. 2022. Slide-Tone and Tilt-Tone: 1-DOF Haptic Techniques for Conveying Shape Characteristics of Graphs to Blind Users. In *CHI Conference on Human Factors in Computing Systems*. 1–19.

[8] Connor Geddes, David R Flatla, Garreth W Tigwell, and Roshan L Peiris. 2022. Improving Colour Patterns to Assist People with Colour Vision Deficiency. In *CHI Conference on Human Factors in Computing Systems*. 1–17.

[9] Andrea Gerino, Lorenzo Picinali, Cristian Bernareggi, Nicolò Alabastro, and Sergio Mascetti. 2015. Towards large scale evaluation of novel sonification techniques for non visual shape exploration. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*. 13–21.

[10] Nicholas A. Giudice, Hari Prasath Palani, Eric Brenner, and Kevin M. Kramer. 2012. Learning Non-Visual Graphical Information Using a Touch-Based Vibro-Audio Interface. In *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility* (Boulder, Colorado, USA) *(ASSETS '12)*. Association for Computing Machinery, New York, NY, USA, 103–110. https://doi.org/10.1145/2384916.2384935

[11] Pedro Antonio Gutiérrez, Maria Perez-Ortiz, Javier Sanchez-Monedero, Francisco Fernandez-Navarro, and Cesar Hervas-Martinez. 2015. Ordinal regression methods: survey and experimental study. *IEEE Transactions on Knowledge and Data Engineering* 28, 1 (2015), 127–146.

[12] Melita Hajdinjak and France Mihelic. 2004. Conducting the Wizard-of-Oz Experiment. *Informatica (Slovenia)* 28, 4 (2004), 425–429.

[13] Highcharts. n.d.. Sonification | Highcharts. https://www.highcharts.com/docs/accessibility/sonification. (Accessed on 08/01/2021).

[14] Leona M Holloway, Cagatay Goncu, Alon Ilsar, Matthew Butler, and Kim Marriott. 2022. Infosonics: Accessible Infographics for People who are Blind using Sonification and Voice. In *CHI Conference on Human Factors in Computing Systems*. 1–13.

[15] Edward Kim and Kathleen F McCoy. 2018. Multimodal deep learning using images and text for information graphic classification. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*. 143–148.

[16] Bongshin Lee, Arjun Srinivasan, Petra Isenberg, John Stasko, et al. 2021. Post-WIMP Interaction for Information Visualization. *Foundations and Trends® in Human-Computer Interaction* 14, 1 (2021), 1–95.

[17] Yotam Mann. n.d.. Tone.js. https://tonejs.github.io/. (Accessed on 08/02/2021).

[18] Kim Marriott, Bongshin Lee, Matthew Butler, Ed Cutrell, Kirsten Ellis, Cagatay Goncu, Marti Hearst, Kathleen McCoy, and Danielle Albers Szafir. 2021. Inclusive data visualization for people with disabilities: a call to action. *Interactions* 28, 3 (2021), 47–51.

[19] David K. McGookin and Stephen A. Brewster. 2006. SoundBar: Exploiting Multiple Views in Multimodal Graph Browsing. In *Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles* (Oslo, Norway) *(NordiCHI '06)*. Association for Computing Machinery, New York, NY, USA, 145–154. https://doi.org/10.1145/1182475.1182491

[20] Giorgio Presti, Dragan Ahmetovic, Mattia Ducci, Cristian Bernareggi, Luca A. Ludovico, Adriano Baratè, Federico Avanzini, and Sergio Mascetti. 2021. Iterative Design of Sonification Techniques to Support People with Visual Impairments in Obstacle Avoidance. *ACM Trans. Access. Comput.* 14, 4, Article 19 (oct 2021), 27 pages. https://doi.org/10.1145/3470649

[21] Ather Sharif, Sanjana Shivani Chintalapati, Jacob O Wobbrock, and Katharina Reinecke. 2021. Understanding Screen-Reader Users' Experiences with Online Data Visualizations. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–16.

[22] Ather Sharif and Babak Forouraghi. 2018. evoGraphs — A jQuery plugin to create web accessible graphs. In *2018 15th IEEE Annual Consumer Communications Networking Conference (CCNC)*. IEEE, Las Vegas, NV, USA, 1–4. https://doi.org/10.1109/CCNC.2018.8319239

[23] Ather Sharif, Olivia H. Wang, Alida T. Muongchan, Katharina Reinecke, and Jacob O. Wobbrock. 2022. Sonifier: JavaScript library that converts a two-dimensional data into a sonified response. https://github.com/athersharif/sonifier. (Accessed on 06/12/2022).

[24] Ather Sharif, Olivia H. Wang, Alida T. Muongchan, Katharina Reinecke, and Jacob O. Wobbrock. 2022. VoxLens: Making Online Data Visualizations Accessible with an Interactive JavaScript Plug-In. In *CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 478, 19 pages. https://doi.org/10.1145/3491102.3517431

[25] Lei Shi, Idan Zelzer, Catherine Feng, and Shiri Azenkot. 2016. Tickers and Talker: An Accessible Labeling Toolkit for 3D Printed Models. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 4896–4907. https://doi.org/10.1145/2858036.2858507

[26] Alexa Siu, Gene SH Kim, Sile O'Modhrain, and Sean Follmer. 2022. Supporting Accessible Data Visualization Through Audio Data Narratives. In *CHI Conference on Human Factors in Computing Systems*. 1–19.

[27] Ruobin Wang, Crescentia Jung, and Yea-Seul Kim. 2022. Seeing Through Sounds: Mapping Auditory Dimensions to Data and Charts for People with Visual Impairments. In *The 24th EG/VGTC Conference on Visualization (EuroVis' 22), Rome, Italy, 13-17 June, 2022*. Eurographics-European Association for Computer Graphics.

[28] Christopher Winship and Robert D Mare. 1984. Regression models with ordinal variables. *American sociological review* (1984), 512–525.

[29] 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.

A PARTICIPANT DEMOGRAPHICS

	Gender	Age	Screen Reader	Vision-Loss Level	Diagnosis
P1	M	57	JAWS	Lost vision gradually	Retinitis Pigmentosa
P2	F	38	JAWS	Blind since birth	Leber Congenital Amaurosis
P3	F	65	JAWS	Lost vision gradually	Retinitis Pigmentosa
P4	F	69	Fusion	Lost vision gradually, Partial vision	Juvenile Macular Degeneration
P5	M	33	NVDA	Blind since birth	Peters Anomaly
P6	M	37	JAWS	Blind since birth	Leber Congenital Amaurosis
P7	F	52	JAWS	Blind since birth	Retinopathy
P8	M	58	JAWS	Lost vision gradually	Cataracts and Glaucoma
P9	M	49	JAWS	Lost vision gradually	Leber Congenital Amaurosis
P10	NB	26	VoiceOver	Partial vision	Corneal damage

Table 1: Screen-reader participants, their gender identification, age, screen reader, vision-loss level, and diagnosis. Under the *Gender* column, M **=** ***Male*****,** F **=** ***Female*****, and** NB **=** ***Non-binary*****.**

B SUBJECTIVE RATINGS

		Average Scores			
Sound Type (*S*)	**Continuous**	*PL*	*CL*	*CF*	*OS*
OmniOscillator with sawtooth waveform (Baseline)	No	3.35	**5.15**	5.10	3.90
	Yes	3.95	4.55	4.70	3.75
OmniOscillator with square waveform	No	**4.65**	5.05	**5.20**	**4.65**
	Yes	4.30	4.50	4.85	4.05
MonoSynthesizer with square waveform	No	3.75	4.95	5.15	4.20
	Yes	3.20	4.15	4.58	3.80

Table 2: Overall average scores for each sound type per continuity level. PL **represents** ***Pleasantness*****,** CL **represents** ***Clarity*****,** CF **represents** ***Confidence*****, and** OS **represents** ***Overall Score*****. Highest average scores for** PL**,** CL**,** CF**, and** OS **are shown in bold.**