

IDENTIFYING A PERCEPTUALLY CONGRUENT FREQUENCY RANGE FOR AUDITORY LINE CHARTS

Joe Fitzpatrick

Dept. of Digital Arts & Media
Technological University of the Shannon
Limerick, Ireland
joe.fitzpatrick@tus.ie

Flaithri Neff

The Software Research Institute
Technological University of the Shannon
Limerick, Ireland
flaithri.neff@tus.ie

ABSTRACT

In this paper, three line chart sonification strategies are contrasted in a user-study examining performance outcomes and perceived difficulty. Auditory line charts are a simple and potentially effective form of parameter-mapping sonification and can be designed easily using freely-available software or a web app. However, an auditory graph's efficacy is affected by a number of design factors such as presentation rate, tonality, and frequency range. Of these three factors, the frequency range remains relatively unexamined in sonification literature and, as such, is the focus of this paper. Three approaches relating to the frequency range are examined, including one that uses a form of amplitude compensation to account for the human frequency response. Results of a user study involving three groups of 16 participants show that a shorter range between 400 Hz and 1400 Hz performs just as well as a wider range between 65 Hz and 1480 Hz. Furthermore, the results suggest that using basic amplitude compensation on the wider range has no significant effect on performance or perceived difficulty.

1. INTRODUCTION

In recent years, new web tools [1, 2], as well as more comprehensive design and evaluation methodologies [3, 4], have helped elevate sonification towards being a viable instrument for relaying scientific data. In relation to user-evaluation, there is an increasing emphasis put on assessing participants' perception of sonic elements within an application's sonification approach [5, 6] as opposed to simply performance-testing it. The heightened interest in examining the perceptual factors at play in applied sonification has produced a spate of design and evaluation frameworks that draw extensively on the wider psychoacoustic knowledgebase [7, 8, 9].

For example, Ferguson and Brewster explore timbral features described as 'roughness' and 'sharpness' in a focus-detection application, but as part of their assessment of application's performance, they examine how closely these timbral features align with the perception of stress, danger, and human error-detection [6]. Their research expands on the idea of basing sound mappings on inherent psychoacoustic sensations that utilize pre-existing cognitive structures extending beyond the auditory periphery [10]. In addition, Ziemer and Schultheis propose using rudimentary visu-

alization and prototypical sonification to benchmark new sonification strategies [8]. This approach contrasts the performance of mapping strategies by placing them on the x and y axes over a mapping area divided into quadrants and fields. Participants can then be tasked with targeting a specific coordinate using various sonifications/visualizations with their performance using each being directly contrastable. Going further, Bonet [5] proposes the establishment of standardized questionnaires to assess the aesthetics of an application as opposed to simply collecting performance and accuracy data.

Perceptual considerations have, to some extent, always been factored into sonification design strategies. Magnitude estimation, for example, has been used to determine what polarity a mapping pair should have (e.g. increasing pitch as temperature goes up being a positive polarity) [11]. Magnitude estimation is a psychophysical method that has been used and explored by researchers to measure the subjective perception of sound intensity and loudness [12, 13]. In such experiments, participants are asked to assign numerical values to the perceived magnitudes of stimuli on a ratio scale without being given any reference stimuli or units of measurement. This approach has been used with reasonable success in a number of sonification studies focussed on auditory perception [11, 14, 15]. This psychoacoustic frame of thinking in relation to sonification is explored in the next section.

2. PSYCHOACOUSTICS IN SONIFICATION

Evaluation methods derived from psychoacoustic paradigms have become increasingly prevalent in sonification design strategies and evaluation techniques. In a recent publication from the Audible Universe project [16], expert opinions were collated with the intention of solidifying design strategies for the sonification of astronomy data. Various direct and indirect approaches for evaluation are identified and disseminated, and include threshold measurements, scaling methods, dissimilarity ratings and more. The result is an informative interdisciplinary dialogue that explores design methodologies and evaluation strategies from multiple unique perspectives. In the exchange, the importance of utilizing psychoacoustic principles is emphasized with evaluation strategies such as threshold measurements being highlighted.

Threshold measurements have been used as a practical means of accounting for perceptual resolution and perceptual interference between auditory dimensions using conventional psychoacoustic JND (just noticeable difference) [9]. In this approach, the measurements are exercised within a x/y orthogonal coordinate system called the 'mapping space', which assists designers in identify-



ing how variable-changes along a mapped dimension might result in perceptual conflicts. This concept uses the maximum likelihood procedure with a number of preformulated psychometric functions that can ultimately infer not only linearity and perceptual interferences, but also on the presence of hysteresis. Designers who engage with the complexity and prerequisites (minor preliminary experiments) of this approach are rewarded with a comprehensive means of evaluating perceptual conflicts during and after the design stages.

The above methodology utilizing JNDs and the maximum likelihood procedure is a good applied example of how psychoacoustic measurements can be incorporated into design pipelines to account for perceptual behavior between two points (or extremes). However, even relatively simple psychoacoustic principles can be used to determine the initial extremes or ranges used in the first place such as the equal-loudness-level contours. Note that Suzuki and Takeshima [17] also utilized the maximum likelihood procedure to determine these contours. Indeed, a good example where the most fundamental of psychoacoustic principles have been effective contributors are in compensation tools such as SuperCollider's AmpCompA - a tool that uses the A-Weighting curve to apply amplitude compensation in real time [18]. The potential benefits of such a tool have also been highlighted in the SoniPy Framework [19], which takes it a step further in proposing a form of 'inverse Fletcher-Munson curve of equal loudness' to counter the variation of perceived loudness at different frequencies. These tools may be useful in exploring various approaches to the design of auditory line charts.

3. AUDITORY LINE CHARTS

Auditory line charts typically map data dimensions (be that single or multiple) to musical notes or tones. Generating auditory graphs has become increasingly easier over the years, initially because of software tools [20, 21, 22, 23], but more recently because of browser-based tools derived from the Web Audio API and various JavaScript libraries such as Tone.js and Sonification.js (Highcharts API) [1, 2, 24, 25]. In the design of auditory line charts, there are a number of factors that affect user interpretation, regardless of dimensionality, such as polarity, tonality, presentation rate, and frequency range. Choices for polarity often depend on the type of data and potentially require a number of preliminary studies to determine. The latter three parameters, however, are also often left as adjustable variables in sonification tools which poses an interesting question as to what guidelines should be suggested when setting them. While the psychoacoustically-driven methodologies highlighted in section 2 are powerful tools for the design and evaluation of individual applications, a broader psychoacoustic framework for the sonification of auditory charts might better prepare practitioners much earlier in the process.

When sonifying a single line trend, tonality is often considered an aesthetic choice as there has yet been no evidence of any significant difference between various instrument-based timbres [26, 27]. Timbre has, however, a practical use when needing to highlight a trend among multiple other trends, and is achievable by using just three to four low-numbered harmonics [28]. This, incidentally, coincides with the IEC 60601-1-8 standard for medical alerts requiring at least four harmonics [29]. Beyond this, there is little evidence to suggest that tonality plays a major role in improving a user's interpretation of a single sonified trend.

In relation to presentation rate of sonified tones, there is still

no conclusive evidence of an 'ideal' rate. In 2003, Brown and Brewster suggested 50-70 ms intervals are preferable [26] however subsequent studies found that there is no significant difference in accuracy between 1 data point per second versus 8 data points per second (125 ms intervals) [30, 31]. In a more recent study, researchers contrasted various presentation rates for sonified EEG trends and discovered that there was no significant difference in accuracy between interval rates of 2ms and 75ms [32], suggesting that presentation rates can be set very close to the interval detection threshold of humans [33, 34]. This lack of a major perceptual significance is reflected in Highcharts and TwoTone's decision to keep presentation rate user adjustable [1, 2]. In summary, presentation rate appears to be considerably flexible without any major negative effect on performance accuracy and can be set according to how the data is scaled.

3.1. Frequency Range

The frequency range is another design factor that is left to the user to determine when using auditory chart tools. While variation in tonality and presentation rate have been shown to have no significant impact on performance outcomes in sonified charts, the same cannot be said of frequency range. Given the considerable perceptual variation demonstrated by the equal-loudness-level contours, it is not unreasonable to assume that frequency-based trend presentations could impact performance. The focus of this paper, therefore, is to examine if variation in frequency ranges affects performance outcomes and perceived difficulty. Also included as part of the study is an examination of loudness compensation to determine if conformity with known-equal-loudness contours is itself an influential factor.

In relation to sonification, early research suggests that designers should not go below MIDI note 35 (61.7Hz) so as to avoid hardware limitations in frequency reproduction or above MIDI note 100 (2637Hz) beyond the centre point of most instrument ranges [27, 26]. Another proposed range in sonification is between 200-5000 Hz where humans are reasonably sensitive and where, again, most instruments' ranges fall [35]. This demonstrates sonification designers' common reliance on music-related factors for setting sonification parameters. This is not illogical as pitch ratios in music reflect our logarithmic perception of frequency resulting from the tonotopic organization of the cochlea and beyond. However, our ability to accurately discern equidistant pitch intervals starts to degrade at higher frequencies. For example, psychoacoustic scales such as the mel scale show discrepancies in perceived pitch intervals above 500Hz [36]. This scale suggests that a person asked to find the approximate half-pitch of 4000 Hz would be more likely to arrive closer to 1050 Hz than 2000 Hz. Subsequent related scales are based on critical bands of the 'auditory filter' concept, whereby overlapping bandpass filters are 'placed' along the basilar membrane [37]. Examples include the Bark scale and various approximations of the equivalent rectangular bandwidth (ERB) number scale, each demonstrating how bandwidth increases significantly as the 'center frequency' increases beyond 500Hz [38, 39]. This is one potential explanation why the Highcharts Sonification Studio has a default (although variable) range of MIDI note 36 (65.4 Hz) to MIDI note 90 (1480Hz). 1500 Hz corresponds to approximately 1300 mels suggesting a relatively low discrepancy in pitch interval perception. 1480 Hz is also the cutoff of the 11th Bark scale critical band after which the bandwidth starts to increase logarithmically.

In addition to interval pitch perception, sonification applications need to account for the varying differences in perceived loudness at different frequencies. It has also been suggested that frequency ranges used in parameter-mapping sonification should be ergonomic and non-fatiguing over time [40, 41]. The equal-loudness-level contours are an authoritative reference for identifying the ‘flattest’ range in the human frequency response curve which roughly corresponds to a 1000 Hz range between 400 Hz to 1400 Hz. Within this subrange, at 60 Phons, frequencies fluctuate by a maximum of 3-4dB SPL. This range also excludes frequencies that we are highly receptive to (3 KHz - 5 kHz) which makes sonified trends less likely to cause fatigue and, lastly, is a range that is easily reproducible. The limitation, however, may be that the range is simply too narrow to adequately reflect the variance of a trend. This range also does not take advantage of our accuracy in identifying equidistant pitch intervals at lower frequencies as mentioned above. One other alternative approach is to use the wider range and account for our uneven frequency response by using an amplitude compensation tool based on the A-weighting curve, such as the SuperCollider tool AmpCompA [18].

4. METHODOLOGY

The aim of the following user study is to compare three sonification approaches that might be used when sonifying single data trends as continuous pure tones:

- 65-1480 Hz (MIDI 36 - MIDI 90) - With no loudness compensation applied.
- 500 - 1400 Hz - Based on the ‘flattest’ range on the equal loudness level contours. Uncompensated.
- 65-1480 Hz (MIDI 36 - MIDI 90) - With loudness compensation applied using AmpCompA[18]

The three possibilities are contrasted in terms of performance (i.e. ability to match sonified trends with visual trends), and perceived difficulty after completing the task.

4.1. Data

The eight data trends that were sonified using the above three approaches were randomly generated to have a normal distribution. They were generated using a mean of 10 and a standard deviation of 2. The eight trends were selected from a wider pool each with a variance ranging from 4.00 - 4.07. The aim here was to limit the overall variance of the sonified trends. In any case, the same eight trends were used for all three approaches, albeit in a randomized order.

4.2. Stimuli

As discussed in section 3.2, the two wider ranges used in approach A and C are more typical in sonification and are loosely based on instrument ranges. In applications and tools that default to these ranges, data variables are often mapped to MIDI notes with changes reflected as discrete note intervals. In this study, however, MIDI was not utilized and instead the note’s approximate respective frequency was used (65-1480 Hz) with trends being sonified as continuous pure tones using Max 8 [42] and SuperCollider. The trends sonified using approach A were done so by mapping the minimum data variable of each to 65 Hz and the maximum to 1480

Hz with everything between scaled linearly. Before being sent to a sine wave oscillator, the variable float ramped continuously between data points at 1 data point-per-second (without any discrete breaks). The result was a continuous uncompensated pure-tone sonification of the trend. For approach B, the same steps were replicated only with the minimum data variable being mapped to 400 Hz and the maximum to 1400 Hz, following that of the ‘flattest’ subrange depicted on the equal-loudness-level contours. Approach C differed in that instead of the scaled data variable going to a Max oscillator, it was instead sent to a SuperCollider session to drive an oscillator with an amplitude-compensated output using the AmpCompA ugen[18]. The result was a continuous pure-tone sonification with a basic frequency-dependent amplitude compensation based on the ANSI-specified A-Weighting curve [43].

4.3. Participants

A total of 48 participants volunteered for this user study with an age range between 20 and 40 years old. All participants were enrolled in a 4-year Music Production and Technology undergraduate programme. As such, at a minimum the participants had a basic understanding of frequency and amplitude as it applies to sound. In order to reduce any mitigating factors such as the potential inconsistent effect of hysteresis on certain trends, or simply some trends being easier to interpret than others, it was necessary to divide the participants into three groups - one for each approach - so that the same trends could be used for each but under different conditions.

4.4. Method

The process the participants went through was the same for all three groups. In a supervised lab, students were given a brief introduction to sonification and auditory line charts before opening a link to the web-based user study. Using their own headphones, the students listened to a preliminary example of a sonified trend while being shown four visual trends. In this preliminary test, participants were shown which trend of the four was the one being sonified. The participants could then proceed with the main study at their own pace and were tasked with identifying which of the four visualized trends were being sonified as seen in Figure 1 below.

There was no time limit set for completing the tests nor were participants required to listen to the entire sonified trend before they could choose their answer. After pairing the sonified trend with one of the four visualized trends, participants were presented with a single ease question (SEQ) in order to assess their perceived difficulty of the task [44]. The question was simply "How difficult was it to pick out the trend?" and required participants to choose from a 7-point Likert (1 - Very Difficult, 7 - Very Easy). An SEQ and Likert was chosen over other more elaborate post-task questionnaires like the Software Usability Scale (SUS) [45] or Subjective Mental Effort Questionnaire (SMEQ) [46] because of its simplicity and general equal performance in sample sizes over 10-12 [47]. Furthermore, the 7-point Likert could be reasonably asked after each of the eight tests so that results could be later aggregated to give a more accurate reflection on overall perceived difficulty as this reduced the chance of the last few tests skewing the result.



Figure 1: Example of one of the tests presented to participants tasked with listening to a sonified trend and choosing its corresponding visual counterpart.

Table 1: The mean, standard deviation and variance of participant accuracy in matching sonified trends to visualized trends. A: 65-1480Hz Uncompensated, B: 400-1400Hz Uncompensated, C: 65-1480Hz Compensated.

Accuracy	A	B	C
Mean	82.0%	90.0%	84.0%
SD	0.21	0.19	0.25
Variance	0.05	0.04	0.06

5. RESULTS

As participants were not presented with any time constraints and could listen to the sonified trends while viewing the four visual trends, it was hypothesized that each group should perform strongly in terms of accuracy. This was indeed the case as the mean accuracy across all 48 participants was 85%. Approach B, which uses the most narrow range of the three performed the strongest with a mean participant accuracy of 90%. However, no significant difference was observed between the three groups as approach A and C were close behind at 82% and 84% respectively. The standard deviation and variation in participant accuracy performance can be seen below in Table 1 and Figure 2.

A single factor ANOVA test contrasting the mean accuracy scores of the three groups determined that, as hypothesized, there was no significant difference between the groups returning a p-value of 0.57. In relation to how difficult the participants of each group found the identifying of sonified trends, a similar result arose with all three groups finding the task moderately difficult with a very low deviation from this verdict as seen in Table 2. Once again, no significant difference was observed in a single factor ANOVA test determining a p-value of 0.64, confirming that all three approaches were perceived equally moderate.

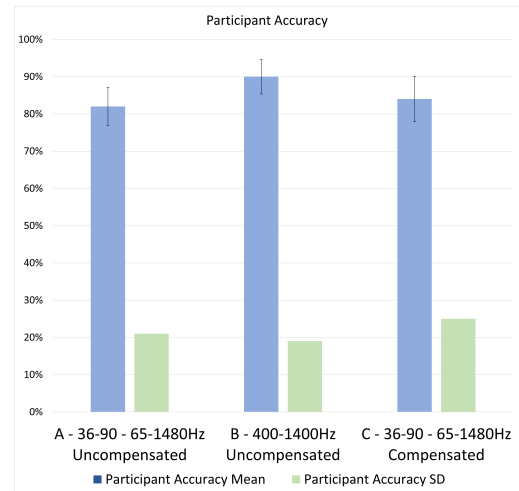


Figure 2: Participant accuracy scores and standard deviation with error bars denoting standard error. Approach C used amplitude compensation.

Table 2: The mean, standard deviation and variance of participants' perceived difficulty in matching trends. A: 65-1480Hz Uncompensated, B: 400-1400Hz Uncompensated, C: 65-1480Hz Compensated.

Perceived Difficulty	A	B	C%
Mean	4.14	4.18	3.89
SD	0.8	0.93	1.1
Variance	0.64	0.86	1.2

6. DISCUSSION

In this study, the performance and perceived difficulty of three sets of sonified trends using three respective approaches has been evaluated. In relation to the stimuli; the data, scaling, presentation rate, and tonality were the same across all three sonification strategies. The design parameter being contrasted in these three approaches was the frequency range. Approach A, as outlined above, uses a frequency range of 65-1480 Hz and is based on MIDI (note 36-90), a relatively wide range where most musical instruments fall within. The main benefit of this range is that it takes advantage of our accuracy in perceiving pitch intervals between 125 Hz and 500 Hz. However, overall perceived loudness of the frequencies between the minimum and maximum of this range vary considerably. This poses the risk of some sections of the sonified trend being perceived as louder than others. Approach C uses the same range, however, the aim here is to improve on performance by using basic amplitude compensation based on the ANSI A-Weighting curve [43]. Approach B uses a different range, 400Hz - 1400 Hz, that is less susceptible to our nonlinear loudness perception of frequency.

The results show that no approach significantly outperformed another, suggesting that approach A was not improved by using amplitude compensation to counteract the human frequency response. It is possible that a more advanced amplitude compensation tool derived directly from the equal-loudness-level contours might be more advisable. However, the equal performance of approach A and C suggests that amplitude compensation is perhaps

unnecessary when sonifying tones continuously. The study has also shown that using a shorter range from 400-1400 Hz is just as viable as a wider range 65 Hz - 1480Hz.

Likert responses also show that there was no significant difference in how difficult each group found interpreting the sonified trends. This outcome is perhaps related to the overall ease of the task presented to them. Another potential means of testing performance might be to present the visual trends after participants have finished listening to the sonified trend. However, such a test is more likely to assess working memory ability, adding an additional variable that may be difficult to account for in the resultant data.

In relation to participant completion times, this may potentially be a usable metric for assessing perceived task difficulty. However, this would only be applicable if participants were asked to try and complete the task as quickly as possible from the outset of the study, which was not the case here. Again, posing such a challenge might be more appreciable in studies examining the auditory glance capabilities of participants. Such a study might also examine the effects of hysteresis on rapidly presented auditory line charts.

To be more congruent with visual line charts, the data points in this study were not sonified as discrete tones but rather as continuous glissando tones that ramp between frequencies. As line charts typically represent continuous data, it may also make sense that sonified trends should be expressed as a continuous wave. An interesting question stemming from this approach then is how prevalent the hysteresis effect is in such sonified charts. For example, while hysteresis has been observed when presenting sequential tones in terms of perceived pitch distance[43], it is unlikely to be a major factor in sonified line charts as the pitch distance between intervals is generally very short (even when data points are discretely sonified subject to data scaling). Chambers et al. [48] found that the perceived direction of a pitch shift is also affected by hysteresis, however, the bias inflicted is a result of the previous trial's relative interval pitch distance as opposed to it simply going 'up' or 'down'. Theoretically, this could affect sonified line charts if they are presented very slowly and the data has a high variance, however, if this is the case line charts are unlikely to have been used in the first place. In any case, continuous tones should limit the risk of hysteresis having an effect on interpreting sonified data-points as the tone ramps between intervals, thus largely negating the hysteresis effect on the perceived pitch distance between each point. A more pertinent question relating to hysteresis and auditory line charts is what the effect of listening to successive sonified trends is. The answer would appear to be once again highly dependent on the overall presentation rate and would be perhaps more specific to applications employing some form of auditory glance.

7. CONCLUSION

In conclusion, this paper demonstrates that no one pitch range outperforms another in the sonification of line chart data. Indeed, a shorter range starting at 400 Hz performs just as well as one starting at 65 Hz. In some circumstances, certain pitch ranges may be purposely chosen by designers for aesthetic reasons or to be reflective of certain chart functionalities - however, these different choices are unlikely to impact on listener performance or on a listener's perception of how difficult the task is. For example, where a playback system has an uneven frequency response or limited low frequency output, an approach similar to that of condition B

in this paper may be more suitable. As another example, the use of smaller frequency ranges might also benefit condensed data trends (or bar charts) by providing a rapid data interpretation consistent with an 'auditory glance'. Although there has been no evidence of hysteresis playing a significant role in the interpretation of sonified line charts, further research might explore its effect on rapidly sonified trends and other forms of auditory charts (bar charts and scatter plots).

Furthermore, additional research on tonality could still improve the interpretability of sonified trends. For example, research to date has primarily focused on comparing instrument-based timbres. Contrasting techniques using noise that are based on well-researched psychoacoustic scales and critical bands might reveal more robust sonification strategies for auditory charts overall. One last dimension that has been so far left unmentioned is the use of space in auditory line charts. Highcharts Sonification Studio, for example, allows designers to use left and right panning to relay the relative position of listeners on the x-axis (for example, the time or date). In addition to tonality, binaural fusion may have an impact on this approach as it relates to our perception of virtual pitch [49].

In sonification, preliminary design considerations that are in line with psychoacoustic principles and implemented from the very beginning should reduce the overall evaluation and performance-testing required. The limitations of human auditory perception, in particular, need to be examined in how they might apply to a final sonified product. In addition, potential improvements, derived from auditory perception research, can only be found if realized and contrasted with typically used sonification strategies. For auditory graphs, this applies to testing the performance when varying adjustable design parameters such as data scaling, presentation rate, tonality, and frequency range; the latter of which has been examined in this paper.

8. REFERENCES

- [1] S. J. Cantrell, B. N. Walker, and Ø. Moseng, "Highcharts sonification studio: an online, open-source, extensible, and accessible data sonification tool." Georgia Institute of Technology, 2021.
- [2] "Twotone sonification tool," <https://twotone.io/>, accessed: February 26, 2023.
- [3] T. Ziemer and H. Schultheis, "Pampas: A psychoacoustical method for the perceptual analysis of multidimensional sonification," 2022.
- [4] G. Parsehian, C. Gondre, M. Aramaki, S. Ystad, and R. Kronland-Martinet, "Comparison and evaluation of sonification strategies for guidance tasks," *IEEE Transactions on Multimedia*, vol. 18, no. 4, pp. 674–686, 2016.
- [5] F. Bonet, "Creating and evaluating aesthetics in sonification," in *Doing Research in Sound Design*. Focal Press, 2021.
- [6] J. Ferguson and S. A. Brewster, "Investigating perceptual congruence between data and display dimensions in sonification," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–9.
- [7] —, "Evaluation of psychoacoustic sound parameters for sonification," in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, 2017, pp. 120–127.
- [8] T. Ziemer and H. Schultheis, "Both rudimentary visualization and prototypical sonification can serve as a benchmark

- to evaluate new sonification designs,” in *Proceedings of the 17th International Audio Mostly Conference, 2022*, pp. 24–31.
- [9] —, “Three orthogonal dimensions for psychoacoustic sonification,” *arXiv preprint arXiv:1912.00766*, 2019.
- [10] J. G. Neuhoff and L. M. Heller, “One small step: Sound sources and events as the basis for auditory graphs,” *Georgia Institute of Technology*, 2005.
- [11] B. N. Walker, “Magnitude estimation of conceptual data dimensions for use in sonification,” *Journal of Experimental Psychology: Applied*, vol. 8, no. 4, pp. 211–221, 2002.
- [12] A. W. Logue, “Individual differences in magnitude estimation of loudness,” *Perception & Psychophysics*, vol. 19, pp. 279–280, 1976.
- [13] M. Epstein and M. Florentine, “Loudness of brief tones measured by magnitude estimation and loudness matching,” *The Journal of the Acoustical Society of America*, vol. 119, no. 4, pp. 1943–1945, 2006.
- [14] B. N. Walker, “Consistency of magnitude estimations with conceptual data dimensions used for sonification,” *Applied Cognitive Psychology*, vol. 21, no. 5, pp. 579–599, 2007.
- [15] J. Ferguson and S. Brewster, “Evaluating the magnitude estimation approach for designing sonification mapping topologies,” *Georgia Institute of Technology*, 2019.
- [16] N. Misdariis, E. Özcan, M. Grassi, S. Pauletto, S. Barras, R. Bresin, and P. Susini, “Sound experts’ perspectives on astronomy sonification projects,” *Nature Astronomy*, pp. 1–7, 2022.
- [17] Y. Suzuki and H. Takeshima, “Equal-loudness-level contours for pure tones,” *The Journal of the Acoustical Society of America*, vol. 116, no. 2, pp. 918–933, 2004.
- [18] SuperCollider, “Ampcompa supercollider help source 3.12,” <https://doc.sccode.org/Classes/AmpCompA.html>, accessed: February 26, 2023.
- [19] D. Worrall and D. Worrall, “The sonipy framework: Getting started,” in *Sonification Design: From Data to Intelligible Soundfields*, 2019, pp. 181–211.
- [20] O. Ben-Tal, J. Berger, B. Cook, M. Daniels, and G. Scavone, “Sonart: The sonification application research toolbox,” *Georgia Institute of Technology*, 2002.
- [21] B. N. Walker and J. T. Cothran, “Sonification sandbox: A graphical toolkit for auditory graphs,” *Georgia Institute of Technology*, 2003.
- [22] B. K. Davison and B. N. Walker, “Sonification sandbox reconstruction: Software standard for auditory graphs,” *Georgia Institute of Technology*, 2007.
- [23] S. Phillips and A. Cabrera, “Sonification workstation,” *Georgia Institute of Technology*, 2019.
- [24] Z. Kondak, T. A. Liang, B. Tomlinson, and B. N. Walker, “Web sonification sandbox-an easy-to-use web application for sonifying data and equations,” *Georgia Institute of Technology*, 2017.
- [25] T. Peng and H. Choi, “Siren: A case study in web audio based sonification,” *Georgia Institute of Technology*, 2021.
- [26] L. M. Brown, S. A. Brewster, S. A. Ramloll, M. Burton, and B. Riedel, “Design guidelines for audio presentation of graphs and tables,” in *International Conference on Auditory Display*, 2003.
- [27] L. M. Brown and S. A. Brewster, “Drawing by ear: Interpreting sonified line graphs,” in *Georgia Institute of Technology*, 2003.
- [28] J. Fitzpatrick and F. Neff, “Stream segregation: Utilizing harmonic variance in auditory graphs,” *Proceedings of SMC*, vol. 18, 2018.
- [29] ISO, “Medical electrical equipment — part 1-8: General requirements for basic safety and essential performance — collateral standard: General requirements, tests and guidance for alarm systems in medical electrical equipment and medical electrical systems,” *International Organization for Standardization*, Geneva, CH, Standard, 2006.
- [30] M. A. Nees and B. N. Walker, “Listener, task, and auditory graph: Toward a conceptual model of auditory graph comprehension,” 2007.
- [31] M. A. Nees, “Data density and trend reversals in auditory graphs: Effects on point estimation and trend identification tasks,” *Master’s thesis*, Georgia Institute of Technology, 2007.
- [32] M. Gavin, R. Jedir, and F. Neff, “Sonification playback rates during matching tasks of visualised and sonified eeg data,” in *Proceedings of the International Audio Systems Symposium*. University of York, 2016.
- [33] C. Muchnik, R. Kaplan-Neeman, N. Aumann, and P. Lavie, “Minimal time interval in auditory temporal resolution,” *J Aud Res*, vol. 25, no. 4, pp. 239–246, 1985.
- [34] K. Krumbholz, E. A. Nobis, A. J. Shahin, W. Nager, C. Braun, T. Weller, M. Schönwiesner, and C. F. Altmann, “Microsecond temporal resolution in monaural hearing without spectral cues?” *J Acoust Soc Am*, vol. 113, no. 5, pp. 2790–2800, 2003.
- [35] B. N. Walker and G. Kramer, “Ecological psychoacoustics and auditory displays: Hearing, grouping, and meaning making,” *Ecological psychoacoustics*, pp. 150–175, 2004.
- [36] J. Volkman, S. S. Stevens, and E. B. Newman, “A scale for the measurement of the psychological magnitude pitch (a),” *J. Acoust. Soc. Am.*, vol. 8, no. 3, pp. 208–208, 1987.
- [37] B. C. Moore and B. R. Glasberg, “Suggested formulae for calculating auditory-filter bandwidths and excitation patterns,” *J. Acoust. Soc. Am.*, vol. 74, pp. 750–753, 1983.
- [38] E. Zwicker, “Subdivision of the audible frequency range into critical bands (frequenzgruppen),” *The Journal of the Acoustical Society of America*, vol. 33, no. 2, pp. 248–248, 1961.
- [39] B. R. Glasberg and B. C. J. Moore, “Derivation of auditory filter shapes from notched-noise data,” *Hearing Research*, vol. 47, no. 1-2, pp. 103–138, 1990.
- [40] T. Hermann, A. Hunt, and J. G. Neuhoff, *The Sonification Handbook*. Berlin: Logos Verlag, 2011, vol. 1.
- [41] G. Parsehian, S. Ystad, M. Aramaki, and R. Kronland-Martinet, “The process of sonification design for guidance tasks,” *J Mob Med*, vol. 9, no. 2, 2015.

- [42] Cycling '74, "Max 8," <https://cycling74.com/>, 2023, accessed: February 26, 2023.
- [43] A. N. S. Institute, "American national standard specification for sound level meters: Ansi s1. 4a-1985 amendment to ansi s1. 4-1983," 1985, acoustical Society of America.
- [44] J. Sauro, "10 things to know about the single ease question (seq)," <https://measuringu.com/seq10/>, Apr. 2012, (Accessed: February 26, 2023).
- [45] J. Brooke, "SUS-a quick and dirty usability scale," *Usability evaluation in industry*, vol. 189, pp. 4–7, 1996.
- [46] E. M. Houwing, M. Weithoff, and A. Arnold, "Cognitive workload measurement. instructions and background materials," TU Delft (WIT Lab), Tech. Rep., 1994.
- [47] J. Sauro and J. S. Dumas, "Comparison of three one-question, post-task usability questionnaires," in *Proceedings of the SIGCHI conference on human factors in computing systems*, Apr. 2009, pp. 1599–1608.
- [48] C. Chambers and D. Pressnitzer, "Perceptual hysteresis in the judgment of auditory pitch shift," *Attention, Perception, & Psychophysics*, vol. 76, pp. 1271–1279, 2014.
- [49] C. Pantev, T. Elbert, B. Ross, C. Eulitz, and E. Terhardt, "Binaural fusion and the representation of virtual pitch in the human auditory cortex," *Hearing Research*, vol. 100, no. 1-2, pp. 164–170, 1996.