RELATIVE INTENSITY OF AUDITORY CONTEXT FOR AUDITORY GRAPH DESIGN

Michael A. Nees and Bruce N. Walker

Sonification Lab, School of Psychology
Georgia Institute of Technology
Atlanta, GA 30332
`mnees@gatech.edu; bruce.walker@psych.gatech.edu`

ABSTRACT

A study examined the role of relative intensity levels for auditory context in auditory graph design. Auditory graphs were designed with auditory context equally as loud as sonified data, context 9 dB more intense than data, or context 9 dB less intense than data. For a point estimation task, participants who experienced auditory graphs with more intense context performed significantly better than participants who experienced graphs with data and context equally loud. Mean differences suggest that making the context either more intense or less intense than the data improved performance as compared to the equally loud condition. We suggest that differences in the intensity of context relative to data facilitate perceptual separation of the auditory streams and thus promote ease of use with auditory graphs. Sound examples are included, and implications for auditory graph design are discussed.

Author Keywords

Auditory graphs, sonification, context, human-computer interaction, sound design

1. INTRODUCTION

Auditory graphs or *sonified graphs* are broadly defined as a class of auditory displays that are produced when sounds are mapped to quantitative data. Auditory graphs commonly use variations in frequency to depict changes in data values (on the visual Y-axis), while the presentations of sounds in time correspond to the visual X-axis. The use of sound to display quantitative information has potential as an alternative or accompaniment to visual data displays, and auditory graphs may prove to be a powerful data analysis tool for people with visual impairments.

Although investigations of auditory graphs remain relatively sparse in the literature, researchers have begun to document their potential. An early investigation of auditory graphs found them to be comparable to the tactile displays that are traditionally used to present graphical information to visually impaired people, and the auditory displays required appreciably less examination time than the tactile displays for some tasks [1]. Flowers and Hauer [2, 3] showed that naïve users could discern information about a data set's distribution from an auditory representation, and Flowers, Buhman, and Turnage [4] demonstrated participants' estimates of Pearson *r* values for bivariate data sets were equivalent regardless of whether the data were presented visually or with sound. Similarly, Bonebright, Nees, Connerley, and McCain [5] found that people could match a bivariate auditory graph to its counterpart visual representation, and Brown and Brewster's [6] participants were able to produce a visual graph from its auditory representation with about 80% accuracy. Not surprisingly, auditory graphs are being implemented in environments where multiple visual displays must be monitored [i.e, financial trading, see 7], and a recent experiment with a divided attention task [8] provided empirical evidence for the viability of auditory graphs in such multi-tasking scenarios.

Despite these promising findings, many questions remain regarding the basic design of auditory representations of quantitative information, and this lack of essential knowledge regarding how exactly to build auditory graphs was a recurring theme at the first International Symposium on Auditory Graphs in 2005 [see 9, 10, 11]. Walker and Nees emphasized the role of auditory context as a means of improving auditory graph design. The data in visual graphs are framed by tick marks, axis labels, legends, etc., and these contextual elements are vital to a viewer's comprehension of the display [12]. Likewise, auditory graph listeners should benefit when sonified data are framed with contextual cues, yet Walker and Nees [11] argued that auditory graph design has, for the most part, been confined to simple tone graphs that are devoid of contextual settings.

Research has suggested that auditory contextual cues can improve performance with auditory graphs. Smith and Walker [13, 14] have shown that sound can be used to provide X and Y axis contextual information that enhances a listener's orientation to the meaning of the actual sonified data. Auditory graphs have been given X axis context in the form of rhythmic beats or clicks [e.g., 5, 13-15], which function as the auditory equivalent of visual tick marks. Furthermore, Y axis context in the form of beeping reference tones have been presented along with the sonified data in order to provide auditory versions of the Y axis gridlines common to visual graphs, and both auditory tick marks and reference tones facilitate performance on point estimation sonification tasks [13, 14].

These findings suggest that design for auditory graphs should incorporate contextual information to orient the listener to the sonified data representations, but basic questions about the implementation of auditory context remain unanswered. Data from a study of training for auditory graphs indicated that some participants had trouble with the perceptual segregation of sonified data and contextual references tones [16]. In particular, the auditory streams representing the data and the context either blended or were confused. Such a confusion can render a sonified graph unusable. Auditory graph research, therefore, should examine design choices that promote perceptual grouping of contextual elements as a separable stream from sonified data [see 17].

One method of promoting the perceptual segregation of data from context in auditory graphs involves altering the relative intensity of data to context in the graph presentation [17].

Intensity changes can be implemented easily in any sound editing program and represent a potential viable approach to facilitating perceptual streaming of respective data and contextual elements. Although researchers have (for good reason) cautioned against using intensity differences to represent changes in a data dimension in auditory graphs [e.g., 9, 18, 19], overall level differences might prove useful for auditory graph designers who wish to promote basic perceptual grouping.

2. METHODS

A study examined the role of relative intensity differences between auditory context and the actual sonified data of an auditory graph. A pilot study was conducted to establish equal perceived loudness settings for data and context tracks within the stimuli. Auditory graph stimuli were then manipulated at three levels: 1) contextual elements of the auditory graph were equally as loud as the data, 2) contextual elements of the graph were 9 dB more intense than the data, or 3) the context was 9 dB less intense than the data. We hypothesized that context presented either softer or louder than the data would promote perceptual segregation of context from data and thus result in better performance on a point estimation task with auditory graphs.

2.1. Stimuli

Auditory graphs represented the price of a stock in dollars over the course of a 10 hr trading day (8 a.m. to 6 p.m.). The price of the stock in dollars (on the Y-axis) was represented by discrete tones that changed in pitch as the price changed, while each hour of the trading day (X-axis) corresponded to one second in time. Discrete tones were presented at the rate of two per second, whereby the pitch of each tone corresponded to the price of the stock at each half-hour of the trading day.

The frequencies of the tones were matched to the price of the stock in dollars using the Sonification Sandbox [20]. The minimum data value ($10) was assigned MIDI note G3, whose frequency is 196 Hz. The maximum data value ($84) was assigned MIDI note B6, whose frequency is 1979.5 Hz. Data were sonified on an exact scale; data values falling between MIDI notes were bent in pitch (rather than rounded to the nearest musical note) to represent the exact frequency of the data point on the scale. A positive polarity was used [21], and stock data were sonified with the piano from the MIDI instrument bank.

Auditory graphs were given Y-axis (price) context using a dynamic Y reference tone, a beeping tone displayed concurrently with the actual discrete data points of the auditory graph. The dynamic Y reference tone created the auditory equivalent of visual gridlines for Y axis values. The tone represented the stock's maximum price of the day ($84, 1979.5 Hz tone) when the true price of the stock was ascending and indicated the stock's lowest price of the day ($10, 196 Hz tone) when the true price was descending. In the current study, maximum value Y context was sonified using an oboe, while minimum value Y context was created with a bassoon. The oboe was chosen because the maximum datum value frequency of 1979.5 Hz falls within the natural range of the oboe, while the minimum datum value frequency of 196 Hz falls within the range of the bassoon.

Time (X-axis) context was provided by the addition of a click track that featured a rhythmic acoustic snare beat every 1 s of the display (i.e. on every hr). The X-axis click track has been shown to be helpful when the density of the track is different than the data density and thus provides added information [13, 14].

2.2. Pilot Study

A pilot study was conducted in order to determine intensity adjustments for auditory context such that listeners perceived the sonified data of the auditory graph stimulus and its respective contextual elements to be equally loud.

MIDI output files from the Sonification Sandbox were converted to .WAV files and imported to Audacity version 1.3.0b. The resulting file had four separate audio tracks: 1) the sonified stock data (piano track), 2) the X axis context (snare drum), 3) the high frequency Y axis context (beeping oboe at 1979.5 Hz), and 4) the low frequency Y axis context (beeping bassoon at 196 Hz). Participants (*N* = 7 Georgia Tech graduate students) were seated facing away from the computer in a room with the experimenter and listened to stimuli through headphones. Participants' task was to adjust each of the three context tracks (snare drum, oboe, and bassoon) to be the same loudness as the sonified data (i.e., the piano track) over the course of the entire auditory graph. Thus, participants were encouraged to make a holistic judgment about the relative loudness of context to data, rather than a local judgment about equal loudness at any specific part of the graph. Participants were allowed to make changes to any of the three context tracks at any time during the procedure, and adjustments were in 3 dB increments. A trial ended when participants indicated that they perceived the four tracks as being equally loud, and the experimenter recorded the intensity adjustments (in dB) required to make each of the respective context tracks sound equally as loud as the sonified data.

Results of the pilot study are shown in Table 1. Participants perceived that all three contextual elements required attenuation in order to sound equally as loud as the sonified data.

Contextual element	Mean Adjustment	SD
X-axis (Snare drum)	- 1.7 dB	2.27
Y-axis (oboe at 1979.5 Hz)	- 16.7 dB	3.05
Y-axis (bassoon at 196 Hz)	- 6.2 dB	3.62

Table 1: Pilot study data. Mean relative adjustments (in dB) where each auditory contextual element was perceived to be equally as loud as the sonified piano data.

Of note, the current study sought to establish and empirically validate some basic guidelines for auditory graph designers regarding relative intensity of data to context in auditory graphs. Many common sound editing programs allow for easy adjustment of the relative intensity of separate audio tracks. Most sound editing programs, however, do *not* allow for separate tracks to be psychophysically matched for equal loudness throughout a recording. We approached the problem, therefore, as a problem of sound design rather than a study in psychophysics.

2.3. Experimental Conditions

A 1 x 3 between-subjects manipulation was employed. The data values in Table 1 were rounded to the nearest 3 dB increment and used as a baseline to manipulate the auditory graph stimuli to create three experimental conditions, as follows.

2.3.1 *Group 1: Data and context equal loudness*

In the equal intensity condition, the relative adjustments in intensity of the snare drum, oboe, and bassoon to the sonified piano data were -3 dB, -18 dB, and -6 dB, respectively, in accordance with the pilot data in Table 1. [ICAD06_graph_equal.wav]

2.3.2 *Group 2: Context lower intensity than data*

The manipulation for this condition was intended to place the contextual elements (snare drum, oboe, and bassoon) of the graph into the background of the auditory listening scene relative to the sonified piano data. Accordingly, the snare drum, oboe, and bassoon were dropped an additional 9 dB lower in intensity than the piano (as compared to the equal intensity condition) to -12 dB, -27 dB, and -15 dB, respectively. The context, therefore, was presented at a lower intensity than the sonified data. [ICAD06_graph_contextsofter.wav]

2.3.3 *Group 3: Context higher intensity than data*

The manipulation for this condition was intended to place the contextual elements of the graph into the auditory foreground relative to the sonified data. The snare drum, oboe, and bassoon were raised 9 dB in intensity (as compared to the equal intensity condition) to +6 dB, -9 dB, and +3 dB, respectively [ICAD06_graph_contextlouder.wav]

2.4. Participants and Design

2.4.1 *Participants*

Participants were 21 Georgia Tech undergraduates (12 females and 9 males, mean age = 20.71 years). All participants reported normal or corrected to normal vision and hearing, and they received course credit for their participation.

2.4.2 *Procedure and Task*

The design of the study featured a pre-test, a brief training session, and a post-test. Participants were randomly assigned to one of the three experimental stimulus conditions described above. Demographic information was collected, and the pre-test session of the study began. Participants received instructions that included a brief description of the auditory graph and task. The pre-test was the same for all participants, and the pre-test auditory graph did *not* employ context. In order to provide a baseline reference, participants were told that the opening price of the stock was 50 dollars. For the 11 pre-test trials, participants were asked to identify the price of the stock for each hour (8 a.m. – 6 p.m.) of the trading day in a randomly selected order. A single trial began with a visual text presentation of the test question (e.g., "What is the price of the stock at 10 a.m.?") followed by the presentation of the auditory graph. Participants were allowed to listen to the auditory graph as many times as needed before responding, and the next trial began after a response had been made using the computer keyboard.

Participants then completed a brief, self-paced training session that gave an overview of auditory graphs, a description of the auditory contextual elements, and a part-task decomposition of the point estimation task. The training session also offered strategies for the successful completion of the task [for a detailed description of the training paradigm, see 14, 15, 16].

The post-test segment of the study began after the training session. All participants were given the same task from the pre-test, where the price of the stock was estimated for each hour of the trading day over 11 trials of randomly selected hours. During the post-test, participants listened to auditory graphs that were enriched with context, and the contextual intensity manipulations described above were employed respectively for the three conditions during the post-test. The dependent variable was defined as the root mean squared (RMS) error (in dollars) of participants' responses to the post-test point estimation trials.

3. RESULTS

Data were analyzed with a one-way Analysis of Covariance (ANCOVA), where pre-test RMS error scores served as a covariate and post-test RMS error scores were the dependent variable. The ANCOVA analysis revealed a significant effect of context group membership, F (2, 17) = 4.37, p = .03. ANCOVA adjusted group means are presented in Figure 1. Tukey's HSD post hoc test revealed a significant mean difference between the context louder group and the context equal group such that performance was significantly better (lower RMS error) when context was louder (mean performance difference = 10.3 dollars, p = .02). Although the difference between the context softer group and the context equal group did not reach statistical significance, the mean difference of 7.16 dollars RMS error suggests a practically relevant finding.

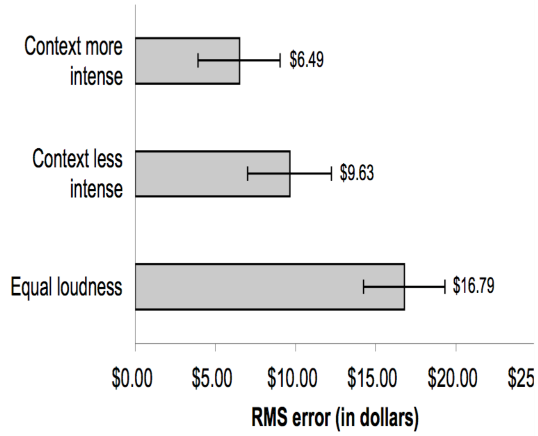

Figure 1. *ANCOVA-adjusted RMS error scores for each of the three context conditions. Note that lower error scores (shorter bars) indicate better performance.*

4. CONCLUSIONS

This study examined the role of the relative intensity of auditory context in the performance of a point estimation task with auditory graphs. Results indicated that auditory graph designs that employ intensity differences for context relative to sonified data resulted in better performance, with 9 dB louder context resulting in the best performance in the current study. The condition where sonified data and auditory context were equally loud resulted in the worst performance. Prior research [16] and theory [17] lead us to explain our results in terms of perceptual streaming. The findings of the current study suggest that altering the intensity of context relative to sonified data in auditory graphs leads to easier perceptual grouping of sounds for data and context, respectively. In other words, participants are better able to use the contextual cues when they are made salient and distinguishable from the data. A visual analogy would be a graph that uses a thick, dark line to represent data and thin, light lines to represent gridlines (or vice versa). Further research with a larger sample size is needed to determine whether reliable statistical and practically relevant differences exist when auditory graphs are designed with louder versus softer context, but our results strongly suggest that sonified graphs should not employ context and data at equal intensities.

As the 2005 International Symposium on Auditory Graphs elucidated, auditory graph researchers have yet to completely identify even the most basic building blocks of design for sonified graphs. The results of this study, however, do offer evidence that designers should consider relative intensity adjustments for auditory context as one way to promote perceptual streaming and build more effective and easier-to-use sonifications.

5. REFERENCES

[1] D. L. Mansur, M. M. Blattner, and K. I. Joy, "Sound graphs: A numerical data analysis method for the blind.," *Journal of Medical Systems*, vol. 9, pp. 163-174, 1985.

[2] J. H. Flowers and T. A. Hauer, "The ear's versus the eye's potential to assess characteristics of numeric data: Are we too visuocentric?," *Behavior Research Methods, Instruments & Computers*, vol. 24, pp. 258-264, 1992.

[3] J. H. Flowers and T. A. Hauer, ""Sound" alternatives to visual graphics for exploratory data analysis.," *Behavior Research Methods, Instruments & Computers*, vol. 25, pp. 242-249, 1993.

[4] J. H. Flowers, D. C. Buhman, and K. D. Turnage, "Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples.," *Human Factors*, vol. 39, pp. 341-351, 1997.

[5] T. L. Bonebright, M. A. Nees, T. T. Connerley, and G. R. McCain, "Testing the effectiveness of sonified graphs for education: A programmatic research project.," presented at International Conference on Auditory Display (ICAD2001), Espoo, Finland, 2001.

[6] L. M. Brown and S. A. Brewster, "Drawing by Ear: Interpreting Sonified Line Graphs," presented at International Conference on Auditory Display (ICAD2003), Boston, MA., 2003.

[7] E. Childs, "Auditory graphs of real-time data," presented at International Conference on Auditory Display (ICAD 2005), Limerick, Ireland, 2005.

[8] S. C. Peres and D. M. Lane, "Auditory graphs: The effects of redundant dimensions and divided attention," presented at International Conference on Auditory Display (ICAD 2005), Limerick, Ireland, 2005.

[9] J. H. Flowers, "Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions," presented at International Conference on Auditory Display (ICAD 2005), Limerick, Ireland, 2005.

[10] T. Stockman, L. V. Nickerson, and G. Hind, "Auditory graphs: A summary of current experience and towards a research agenda," presented at International Conference on Auditory Display (ICAD2005), Limerick, Ireland, 2005.

[11] B. N. Walker and M. A. Nees, "An agenda for research and development of multimodal graphs," presented at International Conference on Auditory Display (ICAD2005), Limerick, Ireland, 2005.

[12] P. A. Carpenter and P. Shah, "A model of the perceptual and conceptual processes in graph comprehension," *Journal of Experimental Psychology: Applied*, vol. 4, pp. 75-100, 1998.

[13] D. R. Smith and B. N. Walker, "Tick-marks, axes, and labels: The effects of adding context to auditory graphs," presented at International Conference on Auditory Display (ICAD2002), Kyoto, Japan, 2002.

[14] D. R. Smith and B. N. Walker, "Effects of auditory context cues and training on performance of a point estimation sonification task," *Journal of Applied Cognitive Psychology*, in press.

[15] B. N. Walker and M. A. Nees, "Brief training for performance of a point estimation task sonification task," presented at International Conference on Auditory Display (ICAD2005), Limerick, Ireland, 2005.

[16] B. N. Walker and M. A. Nees, "Conceptual versus perceptual training for auditory graphs," presented at Human Factors and Ergonomics Society 49th Annual Meeting, Orlando, FL, 2005.

[17] A. S. Bregman, *Auditory scene analysis: The perceptual organization of sound.*, vol. xiii. Cambridge, MA: MIT Press, 1990.

[18] J. G. Neuhoff, G. Kramer, and J. Wayand, "Pitch and loudness interact in auditory displays: Can the data get lost in the map?," *Journal of Experimental Psychology: Applied*, vol. 8, pp. 17-25, 2002.

[19] B. N. Walker and G. Kramer, "Ecological psychoacoustics and auditory displays: Hearing, grouping, and meaning making," in *Ecological psychoacoustics*, J. Neuhoff, Ed. New York: Academic Press, 2004, pp. 150-175.

[20] B. N. Walker and M. Lowey, "Sonification Sandbox: A graphical toolkit for auditory graphs," presented at Rehabilitation Engineering & Assistive Technology Society of America (RESNA) 27th International Conference, Orlando, FL, 2004.

[21] B. N. Walker, "Magnitude estimation of conceptual data dimensions for use in sonification.," in *Journal of Experimental Psychology: Applied*, vol. 8: American Psychological Assn, 2002, pp. 211-221.