PSYCHOPHYSICAL SCALING OF SONIFICATION MAPPINGS: A COMPARISION OF VISUALLY IMPAIRED AND SIGHTED LISTENERS

Bruce N. Walker and David M. Lane

Psychology Department
Rice University
6100 Main St.
Houston, TX 77005 USA
walkerb@rice.edu lane@rice.edu

ABSTRACT

Magnitude estimation was used to determine preferred data-to-display mappings, polarities, and psychophysical scaling functions relating data values (like temperature) to underlying acoustic parameters (like pitch, tempo, or spectral brightness) for blind and visually impaired listeners. The resulting polarities and scaling functions were compared to findings with sighted participants. There was general agreement between the polarities obtained with the two listener populations, with some notable exceptions. There was also evidence for strong similarities with regard to the magnitudes of the slopes of the scaling functions. The results indicate that sonification designers will need to consider whether their intended listeners are visually impaired or not. However, conclusions from this study are limited by the small sample of visually impaired participants. Further research is necessary to arrive at more definitive recommendations.

1. INTRODUCTION

Determining patterns in data is a primary activity for scientists and students. These data sets are increasingly large and complex, making successful scientific exploration an ever-increasing challenge. There are many software tools available for exploring and analyzing data, however they are almost exclusively visual in nature. Such programs do not provide a means for blind and visually impaired students and researchers to participate fully in the scientific endeavor.

Sonification, the use of non-speech audio to display data, can provide crucial data analysis tools for *all* researchers, not only those who are unable to use visual plots and graphs (see [1][2]). However, to ensure that sonification is useful and effective, the auditory display designer must consider the perceptual and cognitive expectancies of the end user the listener and not make design decisions based solely on what sounds "good" or "intuitive" to the designer [2][3]. This may be especially true if the designer happens to be sighted, and the intended listeners are blind or visually impaired.

Walker [2][3][4] points out that to create an effective sonification the designer must determine (1) the optimal display dimension (i.e., sound attribute) to represent the data dimension; (2) the polarity of that mapping; and (3) the scaling of the mapping.

As a concrete example, consider the representation of temperature by the changing frequency of a sound. Within the target group of listeners, perhaps the majority feels that an increase in pitch most obviously represents an increase in temperature. Determining this majority opinion about polarity is the first challenge. Once determined, the designer could use that majority polarity to support design decisions. Next, if the temperature doubled, the designer must know how much to change the frequency in order to represent that temperature change.

The psychophysical paradigm of magnitude estimation [5] (see also [6]) is an effective way to determine both the polarity and the ratio of physical stimulus change to perceived change. The procedure can result in a graph relating the perceived "temperature" to the actual sound frequency. The slope of the line in that graph indicates how much change in frequency is required to represent a given change in temperature. Note that if a doubling of frequency results in a perceived doubling of temperature, then the slope of the graph, or scaling function, would be 1.0. If a doubling of frequency yields less than a doubling in perceived temperature, then the slope of the line would be less than 1.0.

Walker has used magnitude estimation with sighted listeners to answer all three of these questions for several data and display mappings. In addition to charting out the preferred polarities for several data-to-display mappings, Walker [2][4] has found the perhaps surprising result that the actual slope of the scaling function depends on both the sound attribute that is being varied, and the type of data that the sound is supposed to represent. That is, it matters not only how one changes the sound, but also what you call it (such as temperature, velocity, or number of dollars). Participants who were told that some sounds represented pressure yielded slopes that were different from the slopes from participants who heard exactly the same sounds, but were told that they represented temperature. This has significant implications for the design of sonifications, since the actual nature of the data being displayed must be factored in. One size apparently does not fit all.

To date, all of the results in this line of research ([2][3][4]) have been obtained with sighted college students. It is important to continue to replicate and expand the findings in that population. However it is also critical to determine the preferences of other populations, particularly blind and visually impaired listeners. It is not possible to predict in advance if, or how, the mappings, polarities, and scaling functions determined with visually impaired participants might differ from those obtained with sighted students. There are no real theories to predict any differences a priori, although one could postulate differences in the way sound is used to distill information about the environment, or differences in how math and science education affects the perception of data in different populations. Regardless of the actual pattern of results, it is critical to check for any differences, so that sonification design can proceed on a foundation of experimental evidence, rather then speculation.

If the results regarding the preferred polarities and the actual slope values are similar across populations, then development of sonification software may require only one set of synthesis algorithms. However, if different slopes or polarities arise, then auditory display designers and software developers will certainly need to take the broader findings into account. Regardless, the specific needs of visually impaired users must be considered when developing any sonification software.

2. METHOD

This study replicated the procedure used by Walker [2, Experiment 3], but with both sighted and visually impaired participants. Details of the stimuli and experimental procedure are available elsewhere [2][7]. An abridged description is provided here, with departures from the original specified.

2.1. Blind and Visually Impaired Participants

A total of 30 blind and visually impaired youths and adults participated. Fifteen of these participants were adult employees of the Lighthouse of Houston (6 male, 9 female; mean age 37.8 years, range 23-53 years). The other 15 participants were youths from the Texas School for the Blind and Visually Impaired in Austin (11 male, 4 female; mean age 17.5 years, range 12-21 years). All participants were legally blind, though there was a large range in actual visual perception. All participants reported normal hearing, except one male teenager, who had normal hearing in one ear and some hearing loss in the other ear.

2.2. Sighted Undergraduate Participants

Data from the visually impaired participants were compared to data gathered from 83 sighted Rice University undergraduates (22 male, 61 female; mean age 19.7 years, range 18-27 years). All of the sighted participants reported normal or corrected-to-normal vision, and normal hearing.

2.3. Stimuli

This study employed three sets of sound stimuli synthesized in the same way as the sounds used by Walker [2, Experiment 3]. Full synthesis details are provided elsewhere [7], but in brief, the 10 sounds in the Frequency Set were sine tones each 1 s in duration, synthesized at frequencies of 90, 205, 320, 415, 790, 1000, 1350, 1750, 2410, and 3200 Hz. The 10 stimuli in the Tempo Set were each patterns of one beat of sound followed by one-half beat of silence. They were synthesized with a tone frequency of 1000 Hz and were repeated at tempos 41, 60, 107, 167, 203, 270, 415, 505, 572, 685, beats per minute (bpm). The third set, the Brightness Set, was composed of 1-s long FM-synthesized sounds each with a carrier frequency of 100 Hz, a modulation frequency of 300 Hz, and a modulation index (i.e., number of harmonics) of 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10. Increasing the modulation index has the effect of increasing the perceived "brightness" or spectral centroid of the sound. Through pretesting, all sounds within a set were equated for apparent loudness.

Participants made conceptual magnitude estimates of the temperature, pressure, velocity, size, and number of dollars that the sounds seemed to represent.

2.4. Procedure

Each listener participated in three blocks of trials, one for each of the three stimulus sets, with the blocks presented in irregular order. In one block of trials, participants responded to the sounds from the Frequency Set, one sound at a time. In a separate block of trials, participants responded to the Tempo Set. In a third block participants responded to the stimuli in the Brightness Set. The 10 sounds from each of the stimulus sets were presented twice each in random order for a total of 20 trials per block.

The method of modulus-free magnitude estimation was used (see, e.g., [2][5][6]). On each trial, one of the sounds was presented via headphones, and the participant responded with a number that he or she felt estimated the value of the data dimension in use during that block. For example, the participant might listen to sounds of different frequencies, and indicate what "temperature" each sound represented. A sighted assistant helped the visually impaired participants to play the sounds and enter the responses.

3. RESULTS AND DISCUSSION

In all cases, the data from the two experimental groups were analyzed separately. In addition, the two sub-groups within the visually impaired group were first analyzed separately. Although the sample sizes for these sub-groups were too small to do any formal comparisons, an inspection of the data revealed no obvious differences, so all of the data from visually impaired participants were grouped together for subsequent analyses.

First, the data were sorted by display dimension (e.g., frequency), data dimension (e.g., temperature), and participant. For each participant in that block type the Pearson correlation coefficient was calculated between the logarithm of the actual stimulus parameter value (e.g., the frequency) and the logarithm of the values reported (e.g., the perceived temperature). This provided a measure of whether or not a block of data from a given participant exhibited a reliable polarity. If the correlation coefficient did not reach conventional levels of statistical significance, then the data from that participant, in that particular block, were not used in subsequent analyses. Further explanation and justification for this step is provided in [2] and [7].

Next, within each block type, data exhibiting a positive polarity were grouped for analysis separate from data exhibiting a negative polarity.

Then, within each data-to-display pairing and polarity, the data were resorted by stimulus value (e.g., the frequency in Hz), and the geometric mean was calculated for all responses to each individual stimulus, across subjects. For each mapping, the resulting mean data value estimates were plotted against the actual tempos, frequencies, or brightness values of the sounds, on log-log axes. A best-fit line was calculated for each plot, with the slope of the line indicating how much change in, say, temperature was estimated for a given change in the actual frequency of the stimuli.

As an example of the result of these analyses, Figure 1 contains the psychophysical scaling plot for the estimations of temperature for visually impaired listeners; that is, the amount that the perceived temperature changed as a function of the actual frequency change. This plot is representative of the results obtained for all of the data-to-display mappings, though the polarities and actual slopes varied for the different mappings.

Table 1. *Summary of psychophysical scaling slopes with visually impaired listeners*

Display dimension	Slope of regression line (number of participants in that cell)				
	Size	Temperature	Pressure	Velocity	Dollars
Frequency					
Positive Polarity	.71 (2)	.53 (6)	.47 (3)	.90 (4)	
Negative Polarity	-.68 (2)				-.71 (3)
Tempo					
Positive Polarity	.20 (2)	.51 (5)	.72 (6)	.90 (6)	.77 (5)
Negative Polarity					
Brightness					
Positive Polarity	.72 (4)	.42 (4)		.72 (4)	.55 (3)
Negative Polarity					

Table 2. *Summary of psychophysical scaling slopes with sighted listeners*

Display dimension	Slope of regression line (number of participants in that cell)				
	Size	Temperature	Pressure	Velocity	Dollars
Frequency					
Positive Polarity	1.27 (4)	.65 (7)	.79 (4)	.84 (8)	.93 (14)
Negative Polarity	-.56 (5)	-.63 (3)	-.30 (4)		
Tempo					
Positive Polarity		.51 (4)	.64 (6)	.71 (8)	.71 (5)
Negative Polarity	-.74 (6)	-.26 (3)			-.31 (2)
Brightness					
Positive Polarity	.51 (4)	.37 (8)	.69 (6)	.62 (8)	.96 (2)
Negative Polarity	-.18 (4)				-.97 (4)

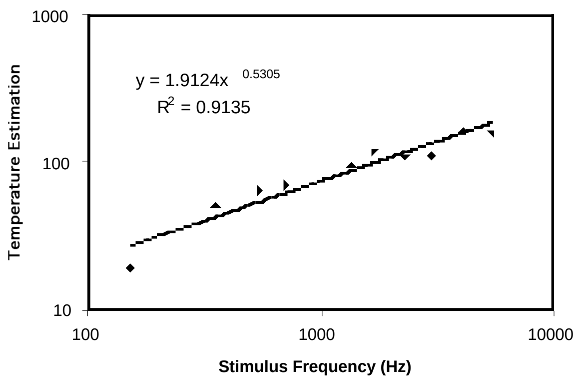

Figure 1. *Temperature estimation versus sound frequency, for visually impaired listeners.*

3.1. Summary of Polarity and Slope Results

Table 1 summarizes the slopes of all of the scaling functions determined in this experiment with blind and visually impaired participants, as well as the number of participants responding with a given polarity (shown in parentheses in the table). Table 2 summarizes the relevant slopes and numbers of participants for the sighted listeners. In both tables, note that a negative slope indicates a negative polarity. That is, an increase in the display dimension (e.g., an increase in frequency) represents a decrease in the data dimension (e.g., a decrease in size). One further note needs to be made about the data that is reported in Tables 1 and 2. The first part of this research is primarily interested in discovering where there is some consensus about polarities and slope values. If there was no consensus about a given mapping polarity (i.e., if fewer than two participants in a given cell responded with a given polarity), those data are not reported here. Since the actual number of participants in each cell was small, due to the overall small pool of visually impaired participants, this results in cells that may be empty here, but will most likely contain data once additional listeners participate in this study, planned for this summer.

3.2. Pattern of Results for Polarity

As was pointed out in the Introduction, it is important to determine first the appropriate polarity of a data-to-sound mapping. This comes primarily from the number of participants who responded to a given mapping with a positive or negative polarity (shown in parentheses in Tables 1 and 2). The polarity with the larger number of participants is considered the majority polarity. Note that there can be ambiguous results and even ties in some cases. While the number of visually impaired participants was only about a third of the number of sighted participants, and therefore limits the conclusions that can be drawn at this point, the results presented in Tables 1 and 2 do have some interesting highlights.

In most cases, the polarity used by the majority of participants for a given data and display dimension pair was the same for both sighted and visually impaired participants. Overall, there was a significant correlation between the number of sighted participants responding with a given polarity and the number of visually impaired participants responding with the same polarity for a given mapping, $r = .38$, $p < .05$. This indicates that in general, there are strong similarities between the preferred polarities shown by sighted and visually impaired listeners.

Notable exceptions include the frequency-dollars, tempo-size, and brightness-dollars mappings. For all three of these mappings, the majority polarity for sighted listeners was opposite the majority polarity obtained for visually impaired listeners. This is particularly striking for the first two cases, where the majority among sighted participants was overwhelming. As discussed, there are no predictive theories about why visually impaired listeners might prefer an opposite polarity. The closest one may come is to offer what seems a plausible explanation, once differences are found. Consider, for example, the frequency-dollars mapping. Sighted participants might be considering that more expensive items within a class, such as automobiles or airplanes, tend to be faster and therefore have higher pitched sounds associated, hence the positive polarity for frequency-dollars. Visually impaired listeners may be more in tune with the everyday sounds of the money itself, noting that a dropped coin makes a high-pitched clink, while a roll of quarters or a bag of notes makes a lower-pitched thud, leading to the inverse polarity for the frequency-dollars mapping. It should be perfectly apparent that any such attempts to explain a mapping are just post hoc rationalizations, and may have absolutely nothing to do with what the listeners are really thinking about. As mentioned, the only reliable way forward is to gather representative data and see what polarities emerge as being preferred.

In the current study, the combination of an overall similarity in response patterns and the presence of some opposite majority polarities underscores the importance of having visually impaired listeners participate in this line of research. It appears that not only the data and display dimensions, but also whether the listener is sighted or not, may need to be factored into any sonifications realistically intended for visually impaired listeners.

3.3. Pattern of Results for Slope

In addition to using the appropriate polarity for a data-to-display mapping, the correct scaling factor needs to be determined to maximize the match between the listener's expectation and the actual sounds presented. It is important to know if visually impaired listeners yield scaling functions (slopes) similar to those obtained with sighted listeners.

Tables 1 and 2 list the exact slope values obtained with visually impaired and sighted listeners, respectively. It is clear that there are differences between the slopes for different data-to-display mappings. This confirms previous results [2][3] that indicate the need to use different scaling functions when designing sonifications that represent different data types.

In addition to examining the specific slopes within a given group of listeners, it is interesting to consider how the overall pattern of responses compares between the two populations. Again, the small sample size for the visually impaired group limits the conclusions that can be drawn here, but these data do contain some interesting findings. Figure 2 compares the absolute values of the slopes obtained in corresponding mappings, for the sighted and visually impaired participants described here. Since the slopes are derived from geometric means computed across the subjects within a mapping type, the means derived from only one or two participants are not very stable. For that reason, and as a compromise due to the small group of visually impaired participants, Figure 2 only presents those slopes that were based on three or more participants' data.

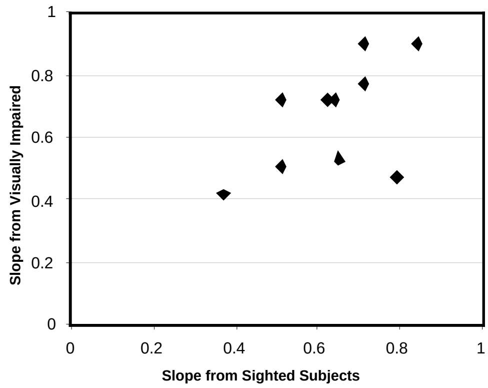

Figure 2. *Correlation between corresponding slopes from sighted and visually impaired listeners.*

With a minimum of three participants per cell, the correlation between the magnitude of the slopes for sighted and visually impaired participant groups is highly significant $r = .82$, $p < .05$. In other words, there is general agreement between the two groups as to how much change is required in a given mapping. This finding with only a 3-participant minimum suggests that once more data are gathered with visually impaired listeners, the slopes obtained for the various data-to-display mappings may be very similar to those reported by sighted participants. Such a result could considerably simplify the process of designing sonifications. Of course, this result only considers the magnitudes of the slopes of the scaling functions, and not their polarities.

4. CONCLUSION

Although these results will need to be replicated and extended with a larger set of participants, the initial implication is that there are many similarities, but some apparently major differences in the way visually impaired and sighted listeners consider sounds to represent data. Simply designing for sighted users will presumably not yield the highest level of comprehension, and therefore effectiveness, of sonifications when used by researchers and students with visual disabilities.

In particular, there seem to be some data-to-display mappings where the majority of visually impaired participants disagree with the polarities preferred by sighted listeners. The exact list, and the nature of these disagreements, needs to be determined in order to apply the appropriate mapping polarity, depending on the target audience for a sonification.

Then, with the appropriate polarity, the correct scaling factor needs to be applied to the mapping. Fortunately, it appears that visually impaired listeners may expect scaling factors that are similar to those expected by sighted listeners. More data need to be gathered before this issue can be resolved, since a limitation of the present study is the relatively small group of visually impaired participants. Continued experimentation in this area should lead quite quickly to effective and valid recommendations for sonifications and auditory displays that will greatly assist both visually impaired and sighted students and scientists.

5. REFERENCES

[1] Kramer, G., Walker, B., Bonebright, T., Cook, P., Flowers, J., Miner, N.; Neuhoff, J., Bargar, R., Barrass, S., Berger, J., Evreinov, G., Fitch, W., Gr hn, M., Handel, S., Kaper, H., Levkowitz, H., Lodha, S., Shinn-Cunningham, B., Simoni, M., Tipei, S. The Sonification Report: Status of the Field and Research Agenda. Report prepared for the National Science Foundation by members of the International Community for Auditory Display. ICAD, Santa Fe, NM, 1999.

[2] Walker, B. N.. *Magnitude estimation of conceptual data dimensions for use in sonification*. Unpublished doctoral dissertation, Rice University, Houston, TX, USA, 2000.

[3] Walker, B. N., & Kramer, G. "Mappings and metaphors in auditory displays: An experimental assessment." In S. Frysinger & G. Kramer (Eds.), *Proceedings of the Third International Conference on Auditory Display, ICAD 96*, (pp. 71-74). ICAD, Palo Alto, CA, 1996.

[4] Walker, B. N., Kramer, G., & Lane, D. M. "Psychophysical scaling of sonification mappings." *Proceedings of the Sixth International Conference on Auditory Display, ICAD 2000*. Atlanta, GA, USA, 2000.

[5] Stevens, S. S. *Psychophysics: Introduction to its perceptual, neural, and social prospects*. Wiley, New York, 1975.

[6] Hellier, E., Edworthy, J., & Dennis, I. "A comparison of different techniques for scaling perceived urgency." *Ergonomics, 38*, 659-670, 1995.

[7] Walker, B. N., & Lane, D. M. "Sonification mappings: A comparison of sighted and visually impaired listeners" Manuscript in preparation.

6. ACKNOWLEDGMENTS

This research was supported by Grant IIS-9906818 from the National Science Foundation.