Universal Design of Auditory Graphs: A Comparison of Sonification Mappings for Visually Impaired and Sighted Listeners

BRUCE N. WALKER and LISA M. MAUNEY
Georgia Institute of Technology

Determining patterns in data is an important and often difficult task for scientists and students. Unfortunately, graphing and analysis software typically is largely inaccessible to users with vision impairment. Using sound to represent data (i.e., sonification or auditory graphs) can make data analysis more accessible; however, there are few guidelines for designing such displays for maximum effectiveness. One crucial yet understudied design issue is exactly how changes in data (e.g., temperature) are mapped onto changes in sound (e.g., pitch), and how this may depend on the specific user. In this study, magnitude estimation was used to determine preferred data-to-display mappings, polarities, and psychophysical scaling functions relating data values to underlying acoustic parameters (frequency, tempo, or modulation index) for blind and visually impaired listeners. The resulting polarities and scaling functions are compared to previous results with sighted participants. There was general agreement about polarities obtained with the two listener populations, with some notable exceptions. There was also evidence for strong similarities regarding the magnitudes of the slopes of the scaling functions, again with some notable differences. For maximum effectiveness, sonification software designers will need to consider carefully their intended users' vision abilities. Practical implications and limitations are discussed.

Categories and Subject Descriptors: H.1.2 [**Models and Principles**]: User/Machine Systems—*Human factors; human information processing*; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems—*Audio input/output*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Auditory (non-speech) feedback; user-centered design*; J.4 [**Computer Application**]: Social and Behavioral Sciences—*Psychology*

General Terms: Experimentation, Human Factors, Measurement, Performance

Additional Key Words and Phrases: Magnitude estimation, visually impaired, auditory display

Initial portions of the data, under preliminary analyses, were reported at the *International Conference on Auditory Display (ICAD'01)*. This article includes additional detail about experimental design, results from additional rounds of evaluation with additional subjects.
Portions of this research were supported by Grant IIS-9906818 from the National Science Foundation.
Authors' addresses: B. N. Walker (corresponding author), L. M. Mauney, Sonification Lab, Georgia Institute of Technology, Atlanta, GA, 30332; email: bruce.walker@psych.gatech.edu.

ACM Reference Format:
Walker, B. N. and Mauney, L. M. 2010. Universal design of auditory graphs: A comparison of sonification mappings for visually impaired and sighted listeners. ACM Trans. Access. Comput. 2, 3, Article 12 (March 2010), 16 pages. DOI = 10.1145/1714458.1714459.
http://doi.acm.org/10.1145/1714458.1714459.

1. INTRODUCTION

Determining patterns in data is a primary activity for scientists and students. As datasets are becoming increasingly large and complex, making successful scientific exploration is 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 nonspeech 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 Kramer et al. [1999], Nees and Walker [2009], Walker and Kramer [2006]). Example domains include the detection of tumors (e.g., Martins and Rangayyan [1997]), mining of stock market data (e.g., Nesbitt and Barrass [2002]), exploration of weather data (e.g., Flowers and Grafel [2002]), interaction with geo-spatial data. [Zhao 2005], and study of engineering data for construction (e.g., Valenzuela et al. [1997]). Encouragingly, work by Brewster [2002] and Brown et al. [2002] demonstrates that blind and vision-impaired individuals can successfully access information from sonified line graphs, both with single [Brewster 2002] and multiple data series [Brown et al. 2002]. Alty and Rigas [1998] also showed that blind users are able to obtain spatial information from graphical shapes through nonspeech sounds (i.e., music). They also found that the context in which the music was presented helped the user obtain more information to better identify the shapes presented in the space. In fact, recognition of the potential utility of sound as a data-display medium has led to software packages that can produce different types of simple "auditory graphs", which are sometimes very useful (even intended) for assisting blind and visually impaired students and scientists. Examples of such software include Triangle [Gardner et al. 1996], the Accessible Graphing Calculator (AGC) [Gardner 1999], the Math Description Engine (MDE) Graphing Calculator [NASA Information Access Lab 2004], the Sonification Sandbox [Davison and Walker 2007; Walker and Cothran 2003; Walker and Lowey 2004], and the vOICe Accessible Graphing Calculator [Meijer 2004].

However, to ensure that sonification and auditory graphs are useful and effective, the auditory display designer must consider, among other issues, the perceptual and cognitive expectancies of the end-user, that is, the listener, and not make design decisions based solely on what sounds "good" or "intuitive" to the designer [Walker 2002; Walker and Kramer 2005; Wickens et al. 2004]. This may be especially true if the designer happens to be sighted, and the intended listeners are blind or visually impaired.

In creating a sonification or an auditory graph, the values in a dataset are often used to vary an acoustic display parameter, such as frequency (pitch), amplitude (loudness), timbre, or tempo, which is intended to represent, or display, the data values. An important issue is the best *mapping* of data values to the available display (sound) dimensions. While frequency is the most commonly used dimension, Walker [2002; 2007] has pointed out that different sound dimensions are better for representing certain data types. For example, in studies with sighted college students, Walker [2002; 2007] and Walker and Kramer [2005] have found that frequency is better for representing temperature, but tempo may be better for size. Next, given a specific mapping, the sonification designer needs to select a polarity and scaling. *Polarity* refers to how the data dimension and the display dimension covary. If a data dimension (e.g., temperature) increases, a positive polarity would dictate that such a change be represented by a corresponding increase in the assigned display dimension (e.g., increasing frequency) (i.e., as temperature rises the frequency of the sound gets higher). A negative polarity would dictate that such a change be represented by a corresponding *decrease* in the assigned display dimension (i.e., as temperature rises the frequency gets lower). *Scaling* refers to how much change in a data dimension is represented by a given change in the display dimension, which is the equivalent of the slope of a line on a visual graph. The "best" scaling value for representing data with sound can depend on the exact type of data and display dimensions in use (e.g., Edworthy et al. [1995; 1991], Hellier et al. [2002], Walker [2002; 2007]). This means that there will be different scaling factors for, say, dollars, temperature, or urgency, when mapped onto frequency. The use of the most preferred parameters should, overall, lead to better performance with an auditory graph or sonification [Neuhoff and Wayand 2002; Walker 2002]. There is no "master" list of such preferred data-to-sound mappings because most of the research done in this field has been with sighted undergraduate students in a limited selection of situations. However, a wider variety of situations and users (e.g., visually impaired, noncollege users, various cultures and ages) should clearly be studied in order to gain a more comprehensive knowledge base that could be used to create such a list.

The psychophysics research paradigm of magnitude estimation (see, e.g., Gescheider [1997], Hellier et al. [1995], Stevens [1975]) is often used to determine the function representing how observers perceive changes in a physical attribute of a stimulus. Magnitude estimation is an effective way to determine both the polarity and the ratio of physical stimulus change to perceived change (the interested reader is referred to Walker [2007] for a discussion of magnitude estimation as it relates to auditory graphs and sonification). The procedure can result in a mathematical function (often a power function) relating the perceived "temperature" to the actual sound frequency. The slope of the line in that function indicates how much change in frequency is required to represent a given change in, say, temperature. If a doubling of frequency results in a perceived doubling of temperature, then the slope of the function, or scaling factor, 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. The polarity comes from the sign of the slope (e.g., a negative slope means a negative polarity).

We have discussed various researchers who demonstrate that sighted and vision-impaired individuals do understand auditory graphs and sonifications. However, as mentioned earlier, the design of auditory displays is often based on decisions of the programmer and not on research. This is where magnitude estimation comes into play; we can use magnitude estimation as a study tool to learn about how people "intuitively" or "naturally" interpret various types of auditory information, which we can, in turn, use to design better auditory displays.

Walker [2002; 2007] has used magnitude estimation with sighted listeners to determine preferred mapping, polarity, and scaling values for several data and display mappings. In addition to charting out the preferred polarities for several data-to-display mappings, Walker [2002; 2007] 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 one calls 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 for participants who heard exactly the same sounds, but were told that they represented "temperature." This has significant implications for the design of sonifications and auditory graphs, since the actual nature of the data being displayed must be factored in. One size apparently does not fit all, both in terms of listeners and datasets.

To date, virtually all of the results in this line of research [Walker 2002; 2007; Walker and Kramer 2005] 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, who are becoming a larger constituency of sonification consumers. 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 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. Empirical results are critical, so that sonification design can proceed on a foundation of scientific evidence.

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. METHODS

This study replicated the procedure that has been used by Walker [2002; 2007], but in this case including blind and visually impaired participants.

2.1 Blind and Visually Impaired Participants

A total of 45 blind and visually impaired youths and adults participated, having been recruited from three organizations. Fifteen participants were adult employees of the Lighthouse of Houston (6 male, 9 female; mean age 37.8 years, range 23–53 years). Another 15 participants were youths from the Texas School for the Blind and Visually Impaired (TSBVI) in Austin (11 male, 4 female; mean age 17.5 years, range 12–21 years). The final 15 participants were employees or clients of the Center for the Visually Impaired (CVI) in Atlanta (6 male, 9 female; mean age 46.6 years, range 28–64 years). This led to a total of 23 male and 22 female participants (overall mean age 34.0 years, range 12–64 years). All participants were legally blind (and so we will use the term "blind" for the remainder of this report to refer to all participants), though there was some variability in actual self-reported visual perception. All participants reported normal hearing, except one male teenager, who reported normal hearing in one ear and some hearing loss in the other ear.

Every participant in this experiment provided signed informed consent, with a sighted assistant reading the consent forms to all of the blind participants, and helping them as needed to sign the consent form. The research protocols were approved by the Institutional Review Board (IRB) of Rice University, of the TSBVI, and of the Georgia Institute of Technology.

2.2 Sighted Undergraduate Data

The data from the blind participants in this study were compared to the extensive dataset obtained by Walker [2007]. For that report, 435 sighted undergraduates with an average age of 20.9 years provided magnitude estimation polarities and slopes using the same stimuli and procedure as was used in the present study. The data from sighted participants were collected in two back-to-back experiments, in a test of the stability of replicating such a magnitude estimation procedure using conceptual data dimensions. To serve as a comparison for the current data, an excerpt of the data from Experiment 1 in Walker [2007] is provided in Table II. Full details are available in the original reports [Walker 2002; 2007].

2.3 Apparatus and Stimuli

The sounds in this study were presented via headphones connected to a computer that displayed the instructions. The headphones used were Sony MDR-V200 (used at TSBVI and the Lighthouse) and Sony MDR-7506 (used at CVI). The computers used to generate the sounds were Apple Macintosh G4 computers with 17-inch studio display monitors (used at TSBVI and the Lighthouse) and an Apple Macintosh Power Book G4 laptop computer (used at CVI). The experiment was written in HTML and JavaScript and ran in Netscape

Navigator 4.6 on Macintosh OS 8.6 (used at TSBVI and the Lighthouse) and in Netscape Navigator 4.7 on Macintosh OS 9.2 (used at CVI).

This study employed three sets of sound stimuli synthesized in 16-bit, 44.1 kHz using Csound to create AIFF files. 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 modulation index 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. All data was saved to text files for later analyses.

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 modulation index 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 standard method of modulus-free magnitude estimation was used (see, e.g., Hellier et al. [1995], Stevens [1975], Walker [2002; 2007]. Since a modulus is a number to which the first stimulus should be associated, the method of modulus-free magnitude estimation allows the participant, and not the experimenter, to set the value for the first stimulus as well as all the remaining stimuli. On each trial, one of the sounds was presented via headphones connected to a computer, and the participant responded to that sound 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. Here is an example of the instructions given to the participants.

> You will hear a series of sounds, one at a time, in random order. Your task is to indicate what size they would represent, by assigning numbers to them. For the first sound, assign it any number of your choosing that represents size. Then, for each of the remaining sounds, estimate its "size", relative to your impression of the first sound. For example, if the second sound seems to represent size that is 10 times that of the first, then assign it a number that is 10 times the first number. If the sound seems to represent size that is

> one-fifth the size of the first, assign it a number that is one-fifth the first number, and so on. You can use any range of numbers, fractions, or decimals that seem appropriate, as long as they are greater than zero.

As indicated in the preceding example instructions, the only range anchors given to participants was that their answers had to be greater than zero.

All participants, sighted and blind, were tested individually. The procedure for the blind and sighted participants differed only in that sighted participants used the computer themselves to play the sounds and enter their own responses; a sighted assistant helped the blind participants interact with the system to perform these tasks.

3. ANALYSES AND RESULTS

3.1 Preanalysis

First, the data from the three subgroups of blind participants (TSBVI, Lighthouse, and CVI) were analyzed separately. A two-way univariate analysis of variance (ANOVA), with an alpha level of 0.05, was used to analyze the data. The ANOVA used the between-subjects factors of group (TSBVI, Lighthouse, CVI) and mapping (e.g., frequency:temperature, tempo:size, etc.), and the dependent variable was ranked slope value. There was no significant main effect of group [$F(2, 17) = 1.332$, $p = .290$] or mapping [$F(14, 17) = .294$, $p = .987$]. There was also no interaction between the two independent variables [$F(27, 17) = .419$, $p = .979$]. This analysis of the data revealed no obvious differences between the groups. Next, the overall combined dataset was split by age of blindness onset into *early* and *late*. Late onset (n = 15) was operationalized as participants who became blind at 18 years of age or later, or participants who had gone through school with vision and thus had some familiarity with visual graphing techniques typically used in schools and society at large. This experience with visual representations of data may cause them to respond differently from early-onset blind participants when asked about how data "should be" represented. The other participants, who became blind before the age of 18, were categorized as early onset (n = 28). It should be noted that it may also have been possible to consider a third subgroup, namely, participants who were blind since birth. However, since our reasons for looking at the early/late-onset distinction was related to experience with visual representations, we did not distinguish in our demographics how many of the early-onset participants were congenitally blind (some were). Regardless, our sample did not contain enough participants in this category to form any viable comparisons. Thus, we maintained the two-group (early/late-onset) distinction. Again, while even these subgroups were too small to make any statistical inferences, an inspection of the results showed the two groups were very similar. Indeed, there was only one case where the early-onset and late-onset groups may have differed. When frequency was used to represent number of dollars, the early-onset blind participants favored a negative polarity, whereas the late-onset participants favored a positive polarity. The possible implications of this are discussed

Table I. Summary of Psychophysical Scaling Slopes with Blind Listeners

Display dimension	Number of participants in each cell (*n*), Slope of regression line									
	Size		Temperature		Pressure		Velocity		Dollars	
Frequency	*n*	*slope*	*n*	*slope*	*n*	*slope*	*n*	*slope*	*n*	*slope*
Positive Polarity	4	0.66	10	0.64	5	0.47	7	0.78	3	1.43
"No" Polarity	0		1		1		1		1	
Negative Polarity	4	–0.62	0	–	2	–0.38	2	–0.23	4	–0.64
Total	8		11		8		10		8	
Tempo										
Positive Polarity	6	0.44	6	0.50	8	0.72	9	0.83	9	0.81
"No" Polarity	1		1		0		0		0	
Negative Polarity	1	–0.83	2	–0.31	2	–0.30	1	–0.68	1	–0.44
Total	8		9		10		10		10	
Modulation Index										
Positive Polarity	6	0.60	6	0.47	4	0.59	7	0.61	5	0.56
"No" Polarity	2		0		1		0		2	
Negative Polarity	1	–0.11	2	–0.43	1	–0.96	2	–0.19	3	–0.43
Total	9		8		6		9		10	

later in the article, along with the other results. Given that there were no systematic subgroup differences in the data, all of the data from blind participants were finally grouped together for the subsequent analyses.

3.2 Individual Analyses of Polarity

Polarity is not typically a critical issue for visual graphs of data, but it can be very important in auditory representations. Within a block, most individual participants apply a consistent mapping polarity (be it positive or negative), and make fairly monotonic responses, so that, for example, low frequencies are given lower numbers and higher frequencies are given higher numbers (or vice versa). In order to separate the positive polarity responses from the negative polarity responses in an algorithmic manner, we used the three polarity categories that have previously been defined: "positive", "negative", and "no" polarity (see Walker [2000; 2002] for more details). For each listener in each block the Pearson correlation coefficient was computed between the log of the responses (e.g., estimated temperature values) and the log of the actual stimulus values (e.g., frequencies). Data from a specific participant in a given block were considered to have "no" polarity, and were not used in subsequent slope analyses, if the absolute value of the correlation coefficient in that block did not reach conventional levels of statistical significance ($r_{critical} = 0.444$, $df = 18$, $\alpha = 0.05$). Stated another way, if the slope of the scaling function was neither positive nor negative, there was "no" polarity. "No" polarity data indicates that either the participant always gave the same or near-same response to whatever sound they heard, or responded in an unsystematic or random manner. These participants either could not distinguish the differences in the sounds, did not understand the experiment, or were not completing the task as instructed. Using this approach, the data in the present study were sorted into positive, negative, and "no" polarity groups for subsequent analyses. These groups are represented in Table I. All 45 participants were included in the analysis, despite the occasional removal of "no" polarity data.

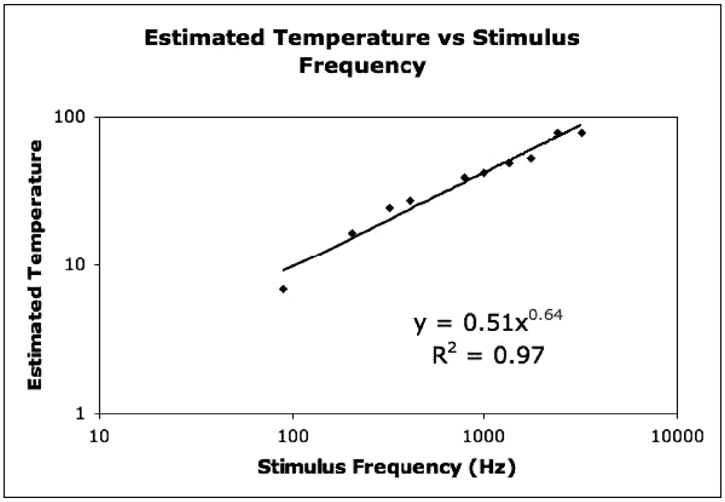

Fig. 1. Temperature estimation versus sound frequency, for visually impaired listeners. The equation of the fit line is shown on the graph. The slope of the fit line represents the scaling factor between frequency of the sounds and the estimated temperature those sounds represent. Note that the positive slope means that this mapping exhibits a positive polarity.

3.3 Aggregate Analyses of Slope

Within each data-to-display pairing and polarity, the data were resorted by stimulus value (e.g., the frequency). Geometric means were calculated for all judgments of a given stimulus, collapsing across participants in a given data and display pair. Geometric means help adjust for different response ranges [Stevens 1975]. These mean estimation values were plotted against the actual stimulus values in log-log coordinates, and fitted with a power function of the form y = b x^m. The exponent, m, which is also the slope of the fit line, indicates how much the perceived or estimated value changes as the actual stimulus parameter changes. As an example of the result of these analyses, Figure 1 contains the psychophysical scaling plot for the estimations of temperature for blind 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 each of the data-to-display mappings, though the polarities and actual slopes varied for the different mappings. For each data-to-display mapping a slope was determined (including both positive and negative slopes, where obtained), along with the number of listeners whose data contributed to each slope. These data are included in Table I.

If both polarities were obtained within a mapping, the majority polarity was defined as greater than 50% of all participants in that block, including the “no” polarity responses (see Walker [2002]). For example, if 6 participants responded with a positive polarity, and 5 responded with a negative polarity, then the positive polarity would be considered as the majority (6 out of 11). However, if there were also two “no” polarities (for a total of 13 participants), that would mean there was no majority for that block. The reasoning behind this

Table II. Summary of Psychophysical Scaling Slopes with Sighted Listeners from Walker [2007, Experiment 1]

Display dimension	Number of participants in each cell (*n*), Slope of regression line									
	Size		Temperature		Pressure		Velocity		Dollars	
Frequency	*n*	*slope*	*n*	*slope*	*n*	*slope*	*n*	*slope*	*n*	*slope*
Positive Polarity	9	0.66	11	0.65	13	0.77	18	0.77	9	1.36
"No" Polarity	0		2		3		1		2	
Negative Polarity	10	–0.87	6	–0.59	3	–0.96	0	–	8	–1.09
Total	19		19		19		19		19	
Tempo										
Positive Polarity	9	0.87	12	0.63	10	0.84	17	0.90	14	1.07
"No" Polarity	1		2		3		1		2	
Negative Polarity	9	–0.77	5	–0.76	6	–0.74	1	–1.29	5	–1.12
Total	19		19		19		19		21	
Modulation Index										
Positive Polarity	9	0.69	14	0.62	17	0.69	17	0.67	12	0.69
"No" Polarity	1		0		0		0		1	
Negative Polarity	9	–0.65	6	–0.48	2	–0.20	2	–0.34	6	–1.03
Total	19		20		19		19		19	

approach is to attempt to determine a "preferred" polarity among the group of participants whenever possible, while at the same time acknowledging that there may not be a unanimous polarity for a given data-to-display pairing. In the practical case of sonification design, an ambiguous or no-majority polarity situation would warrant careful consideration of another mapping, or might indicate the need to explicitly train listeners on how to interpret the display (should such training be effective).

3.4 Summary of Polarity and Slope Results

Table I summarizes the slopes of all of the scaling functions determined in this experiment with blind participants, as well as the number of participants responding with a given polarity. Table II summarizes the relevant slopes and numbers of participants for the sighted listeners from Experiment 1 of Walker [2007]. 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).

3.5 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 respond to a given mapping with a positive or negative polarity. While the number of blind participants was considerably smaller than the number of sighted participants in Walker [2007], which limits the conclusions that can be drawn at this point, the results presented in Tables I and II do have some interesting highlights.

First and foremost, 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 blind participants. Overall, there was a highly significant correlation between the number of sighted participants responding with a given

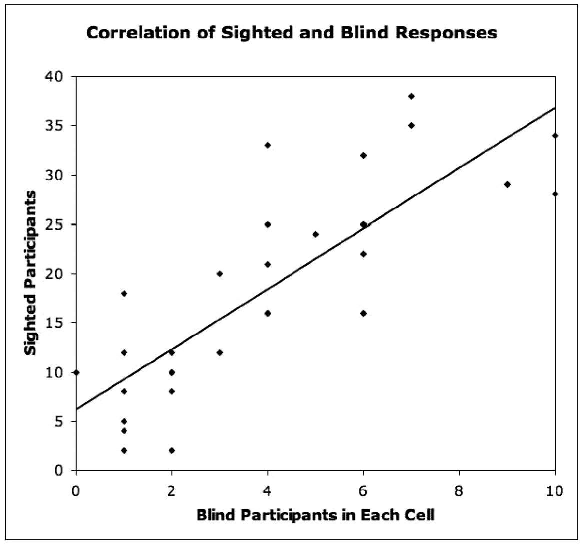

Fig. 2. Correlation between number of respondents in each polarity, for corresponding mappings, from blind listeners in the current study and sighted participants in Walker [2007]. The Pearson correlation between the two series is r = 0.818, p < .001. Note a few points that lie relatively far from the regression line through the data. Those points represent data-to-display mappings where sighted and blind listeners responded somewhat differently. See the text for more discussion.

polarity and the number of blind participants responding with the same polarity for a given mapping ($r = 0.818$, $p < .001$; see Figure 2).

This indicates that in general there are strong similarities between the preferred polarities shown by sighted and blind listeners.

Notable exceptions include the frequency:temperature, tempo:size, tempo: dollars, modulation index:size, and possibly frequency: dollars mappings (compare Tables I and II). For the first four of these mappings, the key difference between the groups was that the blind participants responded with nearly unanimous positive polarities, whereas the sighted participant groups each had both positive and negative polarities represented in significant numbers. This can be contrasted with, for example, the frequency:size mapping where both sighted and blind participants responded with approximately even positive and negative polarities, or the tempo:velocity mapping, where both groups responded with nearly unanimous positive polarity.

For the frequency:dollars mapping in the overall results, the difference in how blind versus sighted listeners interpret the relationship is slight and with additional participants it could change either way. However, interpreting the data that we have, it is perhaps more interesting when the blind participants are divided into early- and late-onset subgroups. As mentioned before, in this mapping early-onset blind listeners preferred the negative polarity (1 positive versus 4 negative), whereas late-onset blind listeners preferred the positive polarity (2 positive versus 0 negative). Thus, it seems that in this particular mapping the late-onset participants respond more like the sighted participants, and less like the early-onset blind participants. This finding is discussed more later.

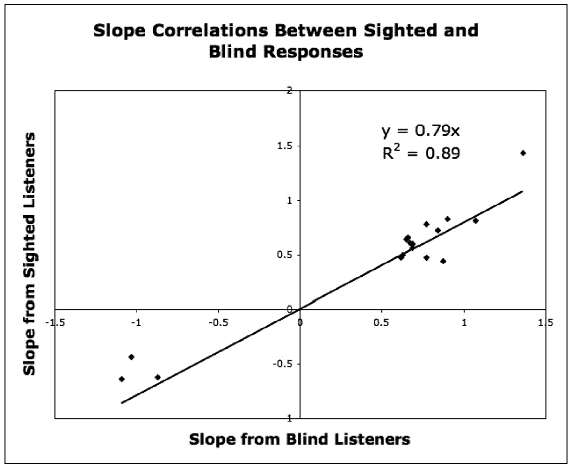

Fig. 3. Correlation between slopes in each polarity, for corresponding mappings, from blind listeners in the current study and sighted participants in Walker [2007]. Slopes were only included in the analysis (and plot) if 3 or more participants contributed values to that cell. The resulting Pearson correlation coefficient between the two series is r = 0.963, p < .001.

3.6 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 in an auditory graph. It is important to know if blind listeners yield scaling functions (slopes) similar to those obtained with sighted listeners.

Tables I and II list the exact slope values obtained with blind and sighted listeners, respectively. It is clear that there are differences between the slopes for different data-to-display mappings. This confirms previous results [Walker 2002; Walker and Kramer 2005] 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 generalizations that can be made here, but these data do contain some interesting findings. Figure 3 compares the slopes obtained in corresponding mappings, for the sighted and blind 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 3 only presents those slopes that were based on three or more participants' data.

With a minimum of three participants per cell, the correlation between the slopes for sighted and visually impaired participant groups is highly significant, $r = 0.963$, $p < .001$. In other words, there is general agreement between the two groups as to how much change (and in which direction) is required in a given mapping. This applies not only to the "central" mappings with slope values in the typical 0.5–0.9 range, but also to the "large slope" mappings such as frequency:dollars, with slopes from both listener groups in the +1.4 range. While the slope values are, overall, highly correlated, the values are often not exactly the same. In most cases this is attributable to uncertainty in the measurement (see Walker [2007]). However, in some cases there is a large difference in slope values that is not likely attributable to measurement error. One such example is the modulation index:dollars (negative polarity) mapping, in which the slope from sighted participants is more than double the corresponding slope from blind listeners.

4. DISCUSSION

The first point to raise is that a psychophysical method often used to gather data from sighted participants (the magnitude estimation procedure) was successfully applied to gather data from blind participants. There were no differences between the three subgroups of blind participants, which allowed us to collapse their data for comparison to sighted listeners. Overall, the data from blind participants were very similar to the data from sighted participants, with a few exceptions. In terms of polarity, sighted participants seem to exhibit more split polarities than blind participants. In the case of the frequency:dollars mapping, given the limits of the data collected here, blind and sighted participants seem to respond with opposite majority polarities.

As discussed, there are no predictive theories about why blind or visually impaired listeners might prefer different polarities than sighted listeners. The closest one may come is to offer what seems a plausible explanation, once differences are found. Consider, for example, the frequency:dollars mapping. Given the difference in responses between sighted and blind listeners (though that requires further study), it would be reasonable to assert that sighted participants likely have never thought about how a somewhat abstract concept like money should be represented by sound. If they have no actual perceptual experience with sounds being related to dollars or money, sighted listeners may rely on their experiences with visual graphs, and may simply resort to a default positive polarity, due to the general usage of positive polarities in visual graphs. This is, in fact, the kind of anecdotal explanation we have often heard from sighted participants. In contrast, visually impaired listeners are likely to be more in tune with the everyday sounds of money itself. Indeed, one blind participant "justified" or explained her responses by noting that a coin dropped on a table makes a high-pitched clink, whereas a roll of quarters makes a clunk, and a bag of coins makes a lower-pitched thud, leading to the negative polarity for the frequency: dollars mapping. This real-life perceptual experience may drive polarity preferences even more for those blind listeners who are not familiar with default visual graphing techniques (i.e., for the early-onset blind participants).

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 (although the previous example is based on comments from a blind participant). As mentioned, the only reliable way forward is to gather representative data and see what preferred polarities emerge.

In the cases where sighted participants show a split polarity, it is important to consider both possible interpretations of a sonification or auditory graph. That is, a display designer may have intended the increasing tempo to represent increasing size, but half the sighted listeners may interpret increasing tempo to mean decreasing size. As a result, the designer needs to know the specific population stereotypes of the intended listeners, and either design to match them, or provide explicit instructions or training to attempt to match listener expectations to the intended message in a display. It is also important to consider when a blind person lost his or her vision, since that may also play a role in interpreting sonifications.

In the current study, the combination of an overall similarity in response patterns and the presence of some differences in preferred 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. Further, it appears that, at least in some cases, late-onset blind participants obtain some experience with visual graphs and, perhaps as a result, may prefer a more visuo-centric, Cartesian-like mapping approach. On the other hand, early-onset blind listeners seem to respond differently, perhaps from not having had the experience with graphs and visual data representations. So, knowing one's users may mean more than just "sighted" or "blind": a listener's specific experience with data representation may play an important role in setting expectations about how data should be represented with sound. Further research is clearly needed in this interesting question.

Once the mapping and polarity are decided upon, the scaling between the display and data dimensions needs to be chosen. The slopes of the magnitude estimation plots are the best starting point at present. The fact that there is very good agreement between sighted and blind participants makes this aspect of the auditory display design task much easier. However, as with polarity, there remain some important differences between the two listening populations, which need to be considered in any real applications. The data in Table I may serve as a starting point.

Although these results will need to be replicated and extended with a larger set of participants, and for a broader selection of data and display dimensions, the initial implication is that there are many similarities, with a few 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 vision disabilities. Continued experimentation in this area should lead quite quickly to even more effective and valid recommendations for sonifications and auditory displays (e.g., Walker and Nees [2008]) that will greatly assist both visually impaired and sighted students and scientists.

REFERENCES

ALTY, J. L. AND RIGAS, D. I. 1998. Communicating graphical information to blind users using music: The role of context. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'98)*. ACM Press/Addison-Wesley Publishing, 574–581.

BREWSTER, S. A. 2002. Visualization tool for blind people using multiple modalities. *Disabil. Rehab. Technol. 24*, 613–621.

BROWN, L., BREWSTER, S. A., RAMLOLL, R., RIEDEL, B., AND YU, W. 2002. Browsing modes for exploring sonified line graphs. In *Proceedings of the British Human-Computer Interface Conference*. 6–9.

DAVISON, B. AND WALKER, B. N. 2007. Sonification Sandbox overhaul: Software standard for auditory graphs. In *Proceedings of the International Conference on Auditory Display (ICAD'07)*. 386–390.

EDWORTHY, J., HELLIER, E. J., AND HARDS, R. 1995. The semantic associations of acoustic parameters commonly used in the design of auditory information and warning signals. *Ergonom. 38*, 2341–2361.

EDWORTHY, J., LOXLEY, S., AND DENNIS, I. 1991. Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Hum. Factors 33*, 205–231.

FLOWERS, J. H. AND GRAFEL, D. C. 2002. Perception of sonified daily weather records. In *Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society (HFES)*.

GARDNER, J. A. 1999. The accessible graphing calculator: A self-voicing graphing scientific calculator for windows. http://dots.physics.orst.edu/calculator/.

GARDNER, J. A., LUNDQUIST, R., AND SAHYUN, S. 1996. TRIANGLE: A practical application of non-speech audio for imparting information. In *Proceedings of the International Conference on Auditory Display*. 59–60.

GESCHEIDER, G. A. 1997. *Psychophysics: The Fundamentals*. Lawrence Erlbaum Associates, Mahwah, NJ.

HELLIER, E. J., EDWORTHY, J., AND DENNIS, I. 1995. A comparison of different techniques for scaling perceived urgency. *Ergonom. 38*, 659–670.

HELLIER, E. J., EDWORTHY, J., WEEDON, B., WALTERS, K., AND ADAMS, A. 2002. The perceived urgency of speech warnings: Semantics versus acoustics. *Hum. Factors 44*, 1–17.

KRAMER, G., WALKER, B. N., BONEBRIGHT, T., COOK, P., FLOWERS, J., MINER, N., NEUHOFF, J. G., BARGAR, R., BARRASS, S., BERGER, J., EVREINOV, G., FITCH, W. T., GRÖHN, M., HANDEL, S., KAPER, H., LEVKOWITZ, H., LODHA, S., SHINN-CUNNINGHAM, B., SIMONI, M., AND TIPEI, S. 1999. The sonification report: Status of the field and research agenda. Rep. prepared for the National Science Foundation by members of the International Community for Auditory Display International Community for Auditory Display (ICAD).

MARTINS, A. C. G. AND RANGAYYAN, R. M. 1997. Experimental evaluation of auditory display and sonification of textured images. In *Proceedings of the 4th International Conference on Auditory Display (ICAD'97)*, E. Mynatt and J. A. Ballas Eds., 129–134.

MEIJER, P. 2004. The voice accessible graphing calculator. http://dots.physics.orst.edu/calculator/.

NASA INFORMATION ACCESS LAB. 2004. NASA LTP information access lab: MDE graphing calculator demonstration (Beta 2.0). http://prime.jsc.nasa.gov/demo/IAL/MDE.html.

NEES, M. A. AND WALKER, B. N. 2009. Auditory interfaces and sonification. In *The Universal Access Handbook*, C. Stephanidis Ed., Lawrence Erlbaum Associates, New York, 507–521.

NESBITT, K. V. AND BARRASS, S. 2002. Evaluation of a multimodal sonification and visualization of depth of market stock data. In *Proceedings of the International Conference on Auditory Display (ICAD'02)*. 233–238.

NEUHOFF, J. G. AND WAYAND, J. 2002. Pitch change, sonification, and musical expertise: Which way is up? In *Proceedings of the International Conference on Auditory Display (ICAD'02)*. R. Nakatsu and H. Kawahara Eds., 351–356.

STEVENS, S. S. 1975. *Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects*. Wiley, New York.

VALENZUELA, M. L., SANSALONE, M. J., KRUMHANSL, C. L., AND STREETT, W. B. 1997. Use of sound for the interpretation of impact-echo signals. In *Proceedings of the 4th International Conference on Auditory Display (ICAD'97)*. E. Mynatt and J. A. Ballas Eds., 47–56.

WALKER, B. N. 2000. Magnitude estimation of conceptual data dimensions for use in sonification. Tech. rep., Psychology Department, Rice University, Houston, TX.

WALKER, B. N. 2002. Magnitude estimation of conceptual data dimensions for use in sonification. *J. Exper. Psych. Appl. 8*, 211–221.

WALKER, B. N. 2007. Consistency of magnitude estimations with conceptual data dimensions used for sonification. *Appl. Cogn. Psych. 21,* 579–599.

WALKER, B. N. AND COTHRAN, J. T. 2003. Sonification Sandbox: A graphical toolkit for auditory graphs. In *Proceedings of the International Conference on Auditory Display (ICAD'03)*. 161–163.

WALKER, B. N. AND KRAMER, G. 2005. Mappings and metaphors in auditory displays: An experimental assessment. *ACM Trans. Appl. Percept. 2*, 407–412.

WALKER, B. N. AND KRAMER, G. 2006. Sonification. In *International Encyclopedia of Ergonomics and Human Factors*, W. Karwowski Ed., CRC Press, New York, 1254–1256.

WALKER, B. N. AND LOWEY, M. 2004. Sonification Sandbox: A graphical toolkit for auditory graphs. In *Proceedings of the Rehabilitation Engineering and Assistive Technology Society of America (RESNA) 27th International Conference*.

WALKER, B. N. AND NEES, M. A. 2008. Data density and trend reversals in auditory graphs: Effects on point estimation and trend identification tasks. *ACM Trans. Appl. Percept. 5*, 1–24.

WICKENS, C. D., LEE, J. D., LIU, Y., AND GORDON-BECKER, S. E. 2004. *An Introduction to Human Factors Engineering*. Pearson Prentice Hall, Upper Saddle River, NJ.

ZHAO, H. 2005. Interactive sonification of geo-referenced data. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'05)*. 1134–1135.

Received November 2008; revised June 2009; accepted June 2009