Sonification Design and Metaphors: Comments on Walker and Kramer, ICAD 1996

BRUCE N. WALKER
Georgia Institute of Technology
and
GREGORY KRAMER
ICAD/Metta Foundation

The original Walker and Kramer paper at ICAD 1996 studied the mapping of data dimensions (e.g., temperature) onto sound dimensions (e.g., pitch). In this commentary we consider the historical and methodological context of that early work, discuss its relevance to the field of auditory display, and point out how it forms part of a body of work with ties to other researchers in the ICAD community.

Categories and Subject Descriptors: A.0 [**General Literature**]: General—*Conference proceedings*; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems—*Audio input/output*; H.5.1 [**Information Interfaces and Presentation**]: User Interfaces—*Auditory (non-speech) feedback*; H.5.5 [**Information Interfaces and Presentation**]: Sound and Music Computing—*Methodologies and techniques*

General Terms: Experimentation, Human Factors, Performance

Additional Key Words and Phrases: Sonification, auditory display, data mapping

1. HISTORICAL CONTEXT

The field of auditory displays and sonification had just been crystallized in the early 1990s following the first ICAD conferences in 1992 and 1994. The publication in 1994 of Gregory Kramer's seminal book [Kramer 1994b] allowed researchers to start to speak a common language, use a shared vocabulary, and belong to a community of like-minded interdisciplinary researchers. However, it was still early days, and the tools, techniques, and even the questions that researchers were considering were vaguely defined at best.

At the time, Bruce Walker was a psychology graduate student with an interest in human–computer interfaces in nontraditional environments. It seemed clear that using sound was an excellent way to communicate information in cases where the operator could not *look at* or could not *see* a visual display. However, it was not at all clear how to actually create, let alone evaluate, such information-rich auditory displays. Gregory Kramer had an overriding interest in general principles of sonification and

Authors' addresses: Bruce N. Walker, Sonification Lab, School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332-0170; email: bruce.walker@psych.gatech.edu; Gregory Kramer, Metta Foundation, Portland, OR, 97229; email: gregory@metta.org.

auditory display design. In particular, he was looking for a research partner to further explore his work on metaphor and affect [Kramer 1994a]. After Albert Bregman introduced us there was an immediate shared interest in applying scientific methods to the questions confronting sonification developers. Together we wanted to investigate empirically the fundamental properties of auditory displays to ensure: (1) the sounds would be heard; (2) they would be interpreted by the listener in the way that the display designer intended; and (3) they would help the listener accomplish the analytic or scientific task for which they were intended. The challenge was that there were no established methods for this kind of investigation. Indeed, many of the questions had never been addressed in any systematic manner. The way forward relied on our combined knowledge, skills, interests, and resources in a truly interdisciplinary approach.

2. RESEARCH PROCESS

We decided that the *mapping* of data onto sound was the place to start [Walker and Kramer 1996]. Sonification mappings, such as using pitch to represent temperature, had often been chosen based on technical considerations, such as the ability to control a certain sound attribute, or based on the designer's intuitions about what sounded "good" or "correct". However, it remained to be determined whether there were more objective, performance-based measures of "goodness" for a data-to-sound mapping. We wondered whether there really was a best way to represent temperature with sound and whether it would be different from representing pressure or velocity, or beauty. In addition, the sound mapping includes the *polarity* of the pairing. That is, whether an increase in pitch best represents an increase or decrease in, say, temperature. In visual graphs, increased distance from the axis typically means increase in the data value. However, with auditory displays there could be other preferred polarities in some cases.

To begin to study mapping and polarity, we developed a simulated production monitoring task. Participants listened to a sonification of the temperature, pressure, size, and rate of production for a fictitious Crystal Factory and made corrective actions when the sound indicated a change in the process. For example, if the sound indicated that the temperature was rising, the corrective action was to turn on a fan. When the pressure was falling, the action was to turn on a pump. We measured both reaction time (RT) and accuracy.

The design of the sonification required us to control four parameters of the sound: pitch, tempo, amplitude onset rate, and loudness. Each sound attribute could also be mapped with a positive or negative polarity. Clearly, to assess all possible combinations of these variables would have been impractical, so we used four sets of mappings (or "ensembles"). The "Intuitive" ensemble was the set of mappings that we (and several other sound designers) felt would be best. Then we created an "Okay" set, and one that we felt would be "Bad" or counterintuitive. A fourth "Random" ensemble was made up of otherwise unused mappings. Finally, we created some variations of these ensembles in which some polarities were flipped. These conditions are clearly not exhaustive, but were intended to provide some insight into performance with different sonification designs.

To actually create the sounds we worked with a mixture of synthesis hardware, software, and Csound code. Typical of the time, the combination of technical, musical, and programming skills required was daunting. Then, we had experiment control software built from scratch in Max by a sound programmer. Participants made responses using a modified MIDI drum machine! Finally, we analyzed the data with psychological statistics software. This experiment truly called on every skill we had and required us to tap into every resource, every contact, every expert we could, just to set it up.

As presented in the Appendix to the original paper, the data for both RT and accuracy showed that the "Bad" and "Random" ensembles actually resulted in better performance than either the "Intuitive" or "Okay" set. This somewhat surprising outcome underscored the need for empirical assessments of sonification designs. The polarity manipulations demonstrated that simply switching the polarity in the "Bad" ensemble made it truly bad (i.e., performance with the "Inverted Bad" ensemble was very slow and inaccurate). Thus, it was clear that mapping and polarity choices were, in some senses, fragile. The specific choices that a designer made might not lead to the best performance and with just one minor tweak the whole thing could go from good to bad.

In the end, we looked at a sort of aggregate of response time and accuracy across a variety of conditions and came up with some mappings that were, on the whole, quite good; some that were clearly not so effective; and some that were either mediocre in terms of performance, or led to different performance depending on the polarity in use. In this study, loudness was good for temperature; pitch was good for rate; and tempo was not very effective overall (presumably because it is at a disadvantage whenever RT is a measure of effectiveness). Onset sharpness was, for the most part, a poor display dimension in our study, but it was surprisingly effective when representing size.

3. BODY OF WORK

The results of this initial experimental assessment led us into a program of investigation in auditory display comprehension. One concern was the issue of static versus dynamic sounds. Walker and Ehrenstein [1997, 2000] considered the effect of dynamic auditory stimuli and whether the perception of changes in the pitch of a sound were independent of the actual pitch of a sound. It turned out that listeners were not able to selectively attend to just the pitch or direction of pitch change. Moreover, pitch information intruded more into responses to pitch change than vice versa. The implication is that auditory display designers can use congruent stimuli (e.g., high pitches getting higher) to help distinguish between high and low pitches; however, if pitch change is the important dimension, designers should restrict the range over which stimulus pitches vary.

In parallel with that research in dynamic sounds, the mappings research continued: it was clear that scaling was also very important. That is, how much change in pitch would be required to represent, say, a doubling of temperature. Walker, Kramer, and Lane [2000] first determined preferred data-to-display mappings by directly asking listeners their opinions and then examined the psychophysical scaling functions relating perceived data values to underlying acoustic parameters. We used magnitude estimation to determine how much a change in a display dimension signified when the listener was thinking about temperature, pressure, velocity, or size. The fact that some participants responded with different polarities underscored the importance of that construct. The actual scaling factors (slopes in the magnitude estimation plots) provided evidence that it matters what sound attribute is in use and what that sound attribute is supposed to represent. This has been replicated on several occasions [Walker 2002] and extended to more sound dimensions and a dozen conceptual data dimensions [Walker 2005].

In addition to the properties of the sounds, the attributes of the listener became a research concern. We first considered musical ability, gender, handedness, and age as individual differences that might affect how sounds are interpreted, but there were no consistent correlations. We did, however, find interesting differences between how sighted and blind listeners respond [Walker and Lane 2001], which we explain with the fact that blind listeners use more physically based mental models to "explain" what they hear. We are continuing this investigation, including different populations [Siebebaler 2003] and, more recently, differences in potentially related cognitive abilities such as working memory capacity [Walker and Mauney 2004].

We have more recently expanded the investigation of auditory display design to the issues of context and training. Context refers to the sounds that do not represent data, but rather contribute to the interpretation of the sounds that do. That is, they serve like the axes and tick marks on a visual graph. It seems that such context can be very beneficial, but only when it adds information, and not just clutter, to the display [Smith and Walker 2002, 2004, in press]. Since most users have little experience with sonifications, training is critical. It remains to be determined how best to provide instruction in sonification comprehension. We have shown that explicit training is better than simple practice and now we are looking at conceptual training versus prompting, practice with feedback, and other approaches [Walker and Nees 2005].

The next major steps in this line of research include more detailed investigations of training, context and design features, and individual differences in perception of auditory display. In addition to this, is an extension into more ecological stimuli, more real-world auditory display tasks, and investigating the role of gestalt formation, affect, and real-time interaction in auditory display design and use [Kramer 2004]. This is in line with a growing trend in perception research toward more ecological psychoacoustics [Neuhoff 2004; Walker and Kramer 2004].

4. RELATIONS TO THE FIELD OF AUDITORY DISPLAY

The importance of deeply understanding the way sounds are used to represent data is important for all types of sonifications. Other researchers have also been studying related issues. Barrass [e.g., 1997] has looked at mapping issues from both a psychological or theoretical perspective as well as a practical need. Hellier and Edworthy [Edworthy et al. 1995; Hellier et al. 1993; Hellier et al. 2002] have looked at the mapping of meaning onto sound, though usually in the context of auditory warnings. Neuhoff has looked at acoustic considerations [Neuhoff and McBeath 1996], multiple mappings [e.g., when two data streams are mapped onto two separate sound attributes, Neuhoff et al. 2002], and individual differences in perception [Neuhoff 1998]. Other investigators have compared visual graphs to auditory graphs [Flowers et al. 1997; Flowers and Hauer 1995] and some have looked at practice and training [Bonebright et al. 2001]. Anderson has studied multidimensional data streams and how to split or combine them for optimal performance [Anderson and Sanderson 2004]. The fundamental questions are being studied more all the time, and certainly not all those studies are represented here. While solid progress is being made, much more remains to be done.

5. CONCLUDING THOUGHTS

Our paper from ICAD 1996 represents a point in time when the field was even newer than it is now. Research directions were open-ended to the point of amorphous and the few of us involved were just putting down whatever orienting markers we could. The papers then were often not as complete as the ones that are now submitted to ICAD. However, that flexibility, the willingness to be on the cutting edge was (and remains, we feel) a hallmark of ICAD. If our paper (and our body of work) has somehow encouraged more scientific evaluation of auditory displays, or if it has provided some reference points in a wilderness of opportunities, then we are pleased and excited to have made such a contribution to the auditory display community.

REFERENCES

ANDERSON, J. AND SANDERSON, P. 2004. Designing sonifications for effective attentional control in complex work domains. In *Proceedings of the 48th Annual Meeting of the Human Factors and Ergonomics Society (HFES2004)*, New Orleans, LA. 1818–1822.

BARRASS, S. 1997. *Auditory Information Design*. Unpublished Dissertation, Australian National University.

BONEBRIGHT, T. L., NEES, M. A., CONNERLEY, T. T., AND MCCAIN, G. R. 2001. Testing the effectiveness of sonified graphs for education: A programmatic research project. In *Proceedings of the International Conference on Auditory Display (ICAD 2001)*. Espoo, Finland. 62–66.

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. *Ergonomics 38*, 11, 2341–2361.

FLOWERS, J. H., BUHMAN, D. C., AND TURNAGE, K. D. 1997. Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples. *Human Factors 39*, 3, 341–351.

FLOWERS, J. H. AND HAUER, T. A. 1995. Musical versus visual graphs: Cross-modal equivalence in perception of time series data. *Human Factors 37*, 3, 553–569.

HELLIER, E., EDWORTHY, J., AND DENNIS, I. 1993. Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. *Human Factors 35,* 693–706.

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

KRAMER, G. 1994a. Some organizing principles for representing data with sound. In *Auditory Display: Sonification, Audification, and Auditory Interfaces*. G. Kramer, Ed. Addison-Wesley, Reading, MA. 185–222.

KRAMER, G., ED. 1994b. *Auditory Display: Sonification, Audification, and Auditory Interfaces*. Addison-Wesley, Reading, MA.

KRAMER, G. 2004. Auditory display as an aid to understanding and empowering human function: Invited Keynote Address at ICAD2004 Sonif(y) Festival. In *Proceedings of the International Conference on Auditory Display (ICAD2004)*, Sydney, Australia.

NEUHOFF, J. G. 1998. Dynamic intensity change influences perceived pitch: Attentional differences between musicians and non-musicians. In *Proceedings of the 16th International Congress on Acoustics and 135th meeting of the Acoustical Society of America*. 2359–2360.

NEUHOFF, J. G., ED. 2004. *Ecological Psychoacoustics*. Academic Press, New York.

NEUHOFF, J. G., KRAMER, G., AND WAYAND, J. 2002. Pitch and loudness interact in auditory displays: Can the data get lost in the map? *Journal of Experimental Psychology: Applied 8*, 1, 17–25.

NEUHOFF, J. G. AND MCBEATH, M. K. 1996. The Doppler Illusion: The influence of dynamic intensity change on perceived pitch. *Journal of Experimental Psychology: Human Perception and Performance 22,* 970–985.

SIEBEBALER, L. M. 2003. *Magnitude Estimation of Sound Attributes Used in Auditory Displays: A Study of Blind and Visually Impaired Listeners*. Unpublished Senior Thesis, Georgia Institute of Technology, Atlanta, GA.

SMITH, D. R. AND WALKER, B. N. 2002. Tick-marks, axes, and labels: The effects of adding context to auditory graphs. In *Proceedings of the International Conference on Auditory Display (ICAD2002)*. Kyoto, Japan. 362–367.

SMITH, D. R. AND WALKER, B. N. 2004. Effects of training and auditory context on performance of a point estimation sonification task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting (HFES2004)*. New Orleans, LA. 1828–1831.

SMITH, D. R. AND WALKER, B. N. in press. Effects of auditory context cues and training on performance of a point estimation sonification task. *Applied Cognitive Psychology*.

WALKER, B. N. 2002. Magnitude estimation of conceptual data dimensions for use in sonification. *Journal of Experimental Psychology: Applied 8*, 4, 211–221.

WALKER, B. N. 2005. Consistency of magnitude estimations with conceptual data dimensions used for sonification. Manuscript submitted for publication.

WALKER, B. N. AND EHRENSTEIN, A. 1997. Congruency effects with dynamic auditory stimuli: Design implications. In *Proceedings of the International Conference on Auditory Display (ICAD1997)*. Palo Alto, CA. 7–12.

WALKER, B. N. AND EHRENSTEIN, A. 2000. Pitch and pitch change interact in auditory displays. *Journal of Experimental Psychology: Applied 6*, 15–30.

WALKER, B. N. AND KRAMER, G. 1996. Mappings and metaphors in auditory displays: An experimental assessment. In *Proceedings of the International Conference on Auditory Display (ICAD1996)*. Palo Alto, CA. 71–74.

WALKER, B. N. AND KRAMER, G. 2004. Ecological psychoacoustics and auditory displays: Hearing, grouping, and meaning making. In *Ecological Psychoacoustics*. J. G. Neuhoff, Ed. Academic Press, New York. 150–175.

WALKER, B. N., KRAMER, G., AND LANE, D. M. 2000. Psychophysical scaling of sonification mappings. In *Proceedings of the International Conference on Auditory Display (ICAD2000)*. Atlanta, GA. 99–102.

WALKER, B. N. AND LANE, D. M. 2001. Psychophysical scaling of sonification mappings: A comparison of visually impaired and sighted listeners. In *Proceedings of the International Conference on Auditory Display (ICAD2001)*. Espoo, Finland. 90–94.

WALKER, B. N. AND MAUNEY, L. M. 2004. Individual differences, cognitive abilities, and the interpretation of auditory graphs. In *Proceedings of the International Conference on Auditory Display (ICAD2004)*. Sydney, Australia.

WALKER, B. N. AND NEES, M. A. 2005. Brief training for performance of a point estimation sonification task. In *Proceedings of the International Conference on Auditory Display (ICAD2005)*. Limerick, Ireland.

Received February 2005; accpeted June 2005