Creating an Empirical Framework for Sonification Design

Janet Anderson

Robens Centre for Health Ergonomics,
European Institute of Health and Medical Sciences,
University of Surrey
Guildford
U.K.
j.anderson@surrey.ac.uk

ABSTRACT

Sonification research and design is held back by a lack of empirical evidence on which to base design decisions. The purpose of this paper is to identify the crucial decisions that need to be made at each stage of the sonification design process and assess what research is required to fill the gaps in the empirical literature. Crucial research questions are identified with the aim of building a framework to guide the decision process.

1. INTRODUCTION

Designers of information displays need to ensure that the display fits the capacities of human operators and supports the tasks that they need to perform. In order to achieve this they need detailed information about how humans perceive and process information and an understanding of the domain in which the display will be used. In visual display design research in these areas is well progressed and there is a large body of evidence for the visual display designer to consult. However, this is not the case for sonification designers. A comprehensive framework is needed to guide sonification designers. The framework should include empirical evidence to guide design decisions at each step of the design process. In this paper I present a potential design framework based on the process of designing a sonification. At each step of the process I highlight the decisions that the designer has to make and the empirical research that is needed to guide those decisions.

Although there is a large literature on auditory perception and attention, this research does not fully address the specific needs of sonification designers. Auditory perception research is often concerned with establishing the performance limits of the human auditory system. Therefore, studies often employ discrete trials, with discrete stimuli that do not change over time. However, the sonification designer needs to know how auditory perception operates with continuous, complex stimuli. As Neuhoff, McBeath & Wanzie [1] have noted, the results of research using static stimuli cannot be generalised to dynamic stimuli. Walker and Kramer [2] also note that the tasks and listening environments used in these studies often bear no relation to the demands of monitoring an auditory display.

In auditory attention research, studies have focused on testing theories of attention, rather than considering the benefits and functions of the appropriate direction of auditory attention [3]. The questions posed by the auditory display designer are more likely to focus on the surface performance of attention than the deep structure [4]. For example, designers need to know how attention operates with multiple inputs, multiple changing sound dimensions and multiple tasks.

Sonification is, by definition, a stream of sound that changes over time and this poses challenges for researchers in devising experimental methods to assess perception and attention. However, a priority is to define the unresolved questions that need to be answered so that sonification designs can be based on firm scientific foundations.

2. EXISTING GUIDELINES

Some general principles for designing auditory displays were discussed by Kramer [5]. These principles, although still untested in a formal sense, can be regarded as an initial attempt to develop guidelines for sonification design. Subsequently, Barrass [6] published a series of five principles, or golden rules for designing auditory displays, and demonstrated how they could be applied. The five principles were directness, appropriateness, range, level and organization.

This detailed work focused on how to analyze the task and the data to ensure that a user's goals are achieved. Although these issues are important, and indeed, were identified by Kramer et al. [7] as pressing issues, they do not answer all the questions that arise in display design. A fine grained analysis of how to configure a sonification is also needed.

3. SONIFICATION DESIGN FRAMEWORK

The design framework presented in this paper is based on a series of studies that investigated the effect of the number of auditory streams, the number of simultaneous changes in the stimulus and the user's attentional set on ability to perceive changes in auditory sequences [8]. The process of designing the stimuli to investigate these questions, extended and expanded to incorporate some extra steps, forms the basis of the present paper. A summary of the framework is shown in Table 1. It includes the design tasks to be completed and for each task the sources of information that will guide the process and the outstanding research questions that need to be answered to inform the task. These are further expanded in the sections below.

3.1. Understand the work domain

Understanding the work domain is crucial to designing effective information displays. Thorough understanding and documentation of the work domain ensures that systems fit the capacities and needs of the domain. Although there are many tools available that analyze various aspects of work domains, in complex socio-technical systems it is important to identify the constraints that exist in the domain so that these can be represented to the operator [9]. Constraints are conditions that must be satisfied in order for the work domain to operate safely and effectively. Cognitive Work Analysis is a suite of methods that can be used to analyze work domain constraints at different levels of the work system [9]. The method is based on the understanding that complex systems often behave in ways that were unanticipated by system designers, and operators are therefore required to reason about the causes of unforeseen system states and correct them. The aim of Cognitive Work Analysis is to identify how that reasoning can be supported. Sanderson, Anderson & Watson [10] have argued that this framework needs to be fully used and extended to provide the information needed for sonification designers.

3.2. Represent higher order relationships

Domain experts often identify problems and make decisions based on their ability to recognize meaningful patterns in the data [11]. Experts use these patterns to form hypotheses about the system state and direct attention to particular data streams to confirm or disconfirm the hypotheses. In support of this contention, previous studies of sonification in anaesthesia have found superior auditory monitoring performance for domain experts compared to student participants [12]. Studying expert performance in the domain will identify the patterns that experts look for and recognize. However, it is not clear how such patterns might be represented in a sonification.

In research on visual displays it has been shown that configural displays, in which an emergent feature (such as the orientation of the line formed by the bars in a bar chart) is mapped to important task variables to indicate system state, can successfully support information integration tasks [13]. Recently, similar results have been found in a study in which a visual object display for anaesthesia was found to support better diagnostic performance [14]. Designing such a display for the auditory modality is a significant challenge. Emergent properties in visual displays include symmetry, closure and parallelism [15], but it is not clear how to represent similar properties in an auditory display [10]. One solution might be to use musical properties such as harmonic resolution or melodic completion to convey higher order information. The effectiveness of incorporating these features into a sonification could be tested empirically using tasks that are ecologically valid for the domain.

3.3. Determine which variables should be displayed aurally

Deciding which variables should be displayed visually and which aurally is a question of major importance that has not received attention in the literature. Although Sanderson et al. [10] suggested that this question could be partly addressed by carrying out a Control Task Analysis (part of Cognitive Work Analysis) it is unclear what criteria would be used to make this decision. Characteristics of the tasks to be performed, the data to be represented and the social environment in which the work is carried out will all affect which variables should be sonified. Research comparing the effectiveness of sonifying different numbers and types of variables versus displaying them visually is needed so that a decision taxonomy can be developed to guide this decision. Although to a certain extent these questions will differ according to the needs of a particular work domain, some general principles could be proposed and tested.

A related question, which appears not to have been addressed in the auditory perception and attention literature, is the number of variables that can be monitored aurally. Patterson, Watts-Perotti & Woods [16] have described how operators use voice loops to communicate and coordinate their activities in air traffic management, aircraft carrier operations and space shuttle mission control. Operators are able to monitor many simultaneous voice loops of communication for relevant information. However, processing of non-speech sound is likely to differ from speech and it is unclear how many data streams listeners can effectively monitor.

In anesthesia sonification research researchers have designed displays of five [17] six [12] and even eight [18] physiological variables. Although the effectiveness of these displays has been tested experimentally, it is difficult to judge the effect of increasing the number of variables in the display because the monitoring tasks used in the various studies were different. A crucial factor that might determine how successfully these high dimensional displays can be monitored is the rate of change in the variables. Human operators calibrate the rate at which they sample visual information displays with the expected rate of change in the signal [4]. However, studies of how people monitor auditory displays have not investigated this issue.

Although some evidence suggests that listeners can detect changes in background sounds that are not in focal attention, depending on the overall stimulus configuration and the attention-attracting qualities of the changes, it is likely that they will need to switch attention to perceive changes accurately [19, 20, 21]. If listeners have to switch auditory attention between multiple auditory inputs and the rate of change in the stimulus is high it is likely that they will miss information.

3.4. Scale the auditory dimensions and the data variables

In order to match the auditory dimensions that will be used to carry information and the data variables, both the auditory dimensions and the data variables need to be scaled. The auditory dimensions need to be divided into discriminable increments so that the range of data values that can be represented can be determined. Similarly, the range of variation to be represented in the data variables should be determined. Careful analysis of operators' tasks will be needed to determine how much information they require about the state of each variable. For example, is it sufficient to know that a variable is outside the normal range, or is information about how far outside normal required? The answer to this question will determine the number of increments needed to represent the variable.

At this stage of the design process, specific information is required about the perceptibility of auditory dimensions and how large changes in the dimensions need to be to be perceptible. Although the psychoacoustic literature has provided valuable data on the sensitivity of the auditory system to changes in auditory dimensions, the specific information required for sonification design is not readily available.

In addition to the artificiality of stimuli and tasks already identified, there are several other factors that limit the applicability of psychoacoustic research to sonification design. Most auditory perception research has investigated only a small number of auditory dimensions. For example, there are relatively large numbers of studies investigating frequency and amplitude, and fewer that investigate other acoustic dimensions such as duration or vibrato. Kramer [5] gave detailed advice about possible dimensions to use for auditory display and recommended that nesting of loudness, pitch and brightness parameters might be a way to increase the number of acoustic dimensions available. He also identified reverberation time, flange depth, speed, wave shape and resonance as potential sonification dimensions. However, there is little information about the discriminability of these dimensions within the range that would be suitable for a sonification.

Another factor limiting the relevance of psychoacoustic research is the fact that the fine discriminations required to detect changes of the magnitude of "just noticeable differences" would not be effective in applied settings, where changes in the sound are required to draw the listener's attention away from other tasks. Studies that investigate how big a change in a dimension is required to attract attention when a listener is engaged in another task are required.

Finally, in monitoring a sonification, operators need to do more than simply detect a change in the sound. They often need to know whether the values of the data stream are increasing or decreasing. Producing a discriminable change in a sound is much easier than the challenges of producing a linearly ordered dimension. Anderson [8] used timbre to produce a linearly ordered dimension that clearly moved from one pole to another. More information is needed on how this might be achieved with other dimensions. At present the effect of the direction of the change in a dimension on people's ability to perceive a change is poorly understood. Neuhoff et al. [1, 22] have found that the interaction between pitch and loudness depends on the relative direction of the changes in each dimension and it is important to extend this research to include other auditory dimensions.

3.5. Map data to sound parameters

The decision about which data variable should be mapped to which sound parameter will be partly determined by the preceding steps which are in effect an analysis of the requirements for the data to be usable and for the sound to be perceptible. The preceding steps will have revealed factors that constrain the choices that can be made at this stage.

Another consideration at this stage of the design process is the degree to which perceptual interactions will affect the intelligibility of the sonification. When monitoring a complex auditory display with many variables, it is likely that a number of variables might change at the same time. It is therefore important to understand how sounds that vary on several dimensions at once are processed. Are listeners able to attend to one changing dimension independently of changes in another dimension?

The evidence suggests that processing can be both enhanced and degraded by simultaneous changes in more than one auditory dimension. For example, pitch and rhythmic structure interact, with a rhythmically accented pitch being recognized more accurately than one that is not accented [23].

In auditory classification tasks using static stimuli, perceptual interactions have been found between pitch and timbre, timbre and loudness, pitch and loudness [24, 25, 26] and location and frequency [27]. These findings indicate that it takes time to filter out information about one of the dimensions and selectively attend to the other dimension. Perceptual interactions between pitch and loudness have also been found with dynamic stimuli [1, 22].

It is not clear from the literature, however, to what extent performance would be reduced by these problems and whether the sonification could be configured to reduce interactions between dimensions. There is evidence that sonification configuration can affect perception, but the mechanisms involved are not clear. For example, Anderson [8] found that perceptual interactions could be reduced by mapping interacting dimensions to different auditory streams and that a change in sonification configuration produced a facilitation effect between some concurrently changing dimensions.

Although some perceptual interactions are extensively documented (for example, equal loudness contours) and appear to arise from non-linearity in the auditory system, other instances of reduced monitoring accuracy when multiple dimensions change might arise from cognitive factors such as attention. Perceptually salient dimensions might draw attention away from less salient dimensions. Dimensions that are based on similar percepts such as the temporal properties of the sound might be highly confusable. These cognitive factors might be amenable to training in strategies for monitoring sounds. The effect of training on people's ability to accurately discriminate changes in an auditory dimension when other dimensions also change appears not to have been investigated.

In summary, sonification designers need empirical evidence to guide them in what strategies increase intelligibility when there are multiple sound dimensions changing. Research is needed on the nature and causes of interactions between dimensions and needs to include a wide range of dimensions. Studies examining the effect of different sonification configurations on interactions between dimensions are needed and the potential of training strategies to reduce difficulties in monitoring multiple changing dimensions needs to be evaluated.

3.6. Determine the number of auditory streams

The impact of the number of auditory streams in a sonification has not been fully investigated. Although monitoring accuracy decreases as the number of auditory streams increases, the relationship depends on the number of dimensions that change and whether attention is selective or divided [8]. These findings need to be extended using different combinations of dimensions, and different tasks.

In a multiple stream sonification the streams have to be distinguished from each other to aid selective attention to a particular stream. This is usually achieved by having a separate pitch range, distinctive timbre or different speeds for each of the streams. Although research in Auditory Scene Analysis [28] has extensively investigated the factors that give rise to the streaming phenomenon, it is not clear how applicable these findings are to sonification. The relative impact of having different pitches, timbres and speeds in creating separate perceptual streams needs to be investigated. These questions are important because the need to separate streams on pitch or timbre or speed then constrains the range of those dimensions on each stream and affects the data variables that can be mapped to each stream.

3.7. Map dimensions to streams

Dimensions can be mapped to streams in two ways: repeated or unique. A repeated mapping means that the same auditory dimensions are used to carry information on each stream. For example, if pitch and speed change on each of two streams, this is a repeated mapping. A unique mapping uses different auditory dimensions to carry information on each stream. This is a potential way to aid attention allocation to individual streams. If pitch only ever changes on stream 1, never on stream 2, attention is cued to stream 1 when a pitch change occurs. However, a unique mapping also means that for some streams information will be carried on less perceptually salient dimensions than other streams. In a repeated mapping of dimensions to streams the most discriminable dimensions can be used for each stream. This also carries the attendant risk that listeners will find it difficult to discriminate streams and might become confused about which of the two or three data variables represented by pitch they heard change.

These questions all need to be addressed in experimental research that assesses the potential advantages and disadvantages of each mapping.

4. CONCLUSION

This paper has identified some future research directions that are crucial for the development of sonification research. The research questions were identified by following a sonification design process and examining the decisions that a sonification designer would have to make at each step. In reality the design process is iterative rather than linear and decisions taken at one stage of the process are likely to constrain later decisions. Nevertheless, the aim was to use the design process as an aid to identify unanswered questions and gaps in the research literature that need to be filled.

5. REFERENCES

[1] Neuhoff, J., McBeath, M. & Wanzie, W. Dynamic frequency change influences loudness perception: A central, analytic process. Journal of Experimental Psychology: Human Perception and Performance, Vol. 25, pp. 1050-1059, 1999.

[2] Walker, B & Kramer, G. Ecological psychoacoustics and auditory display: Hearing, grouping and meaning making. In Neuhoff, J. (Ed.), Ecological psychoacoustics. New York: Academic Press. 2004.

[3] Moray, N. Attention to dynamic visual displays in man-machine systems. In Parasuraman, R. & Davies, D.R. (Eds.), Varieties of attention. (pp. 485-512). Orlando, Florida: Academic Press. 1984.

[4] Moray, N. Designing for attention. In Baddeley, A. & Weiskrantz, L. (Eds.), Attention: Selection, awareness and control: A tribute to Donald Broadbent. (pp. 111-134). Oxford: Clarendon Press. 1993.

[5] Kramer, G. Some organising principles for representing data with sound. In G. Kramer (Ed.), Auditory display: Sonification, audification, and auditory interfaces (pp. 79-94). Reading, MA: Addison-Wesley. 1994.

[6] Barrass, S. Some golden rules for designing auditory displays. In B. Vercoe and R. Boulanger (Eds.), The Csound book. Cambridge MA: MIT Press. 1997.

[7] Kramer, G., Walker, B., Bonebright, T., Cook, P., Flower, J., Miner, N., Neuhoff, J., Bargar, R., Barrass, S., Berger, J., Evreinov, G., Fitch, W.T., Grohn, 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. Santa Fe NM: ICAD. 1999.

[8] Anderson, J. Sonification design for complex work domains: Streams, mappings and attention. Unpublished PhD thesis, School of Psychology, The University of Queensland, St. Lucia, Qld. 2004.

[9] Vicente, K. Cognitive Work Analysis. Mahwah, NJ: Erlbaum. 1999.

[10] Sanderson, P. M., Anderson, J. & Watson, M. Extending ecological interface design to auditory displays. Proceedings of the 2000 Annual Conference of the Computer-Human Interaction Special Interest Group (CHISIG) of the Ergonomics Society of Australia (OzCHI2000). Sydney, Australia. December 4-8, 2000.

[11] Klein, G. A recognition-primed decision (RPD) model of rapid decision making. In G. Klein, J. Orasanu, R. Calderwood & C. Azambok (Eds.), Decision making in action: Models and methods. Norwood, NJ: Ablen Publishing Corp. 1993.

[12] Watson, M. & Sanderson Sonification helps eyes-free respiratory monitoring and task time-sharing. Human Factors, Vol. 46, pp. 2004.

[13] Sanderson, P., Flach, J., Buttiegieg, M. & Casey, E. Object displays do not always support better integrated task performance. Human Factors, Vol. 31, pp. 183-198, 1989.

[14] Blike, G., Surgenor, & Whalen, K. A graphical object display improves anesthesiologists' performance on a simulated diagnostic task. Journal of Clinical Monitoring and Computing, Vol. 15, pp. 37-44, 1999.

[15] Bennett, K., Nagy, A. & Flach, J. Visual displays. In G. Salvendy (Ed.), Handbook of human factors and ergonomics (pp. 659-696). New York: John Wiley. 1997.

[16] Patterson, E.S., Watts-Perotti, J. & Woods, D.D. Voice loops as coordination aids in space shuttle mission control. Computer Supported Cooperative Work, Vol. 8, pp. 353-371, 1999.

[17] Loeb, R.G., & Fitch, W.T. A laboratory evaluation of an auditory display designed to enhance intra-operative monitoring. Anesthesia and Analgesia, Vol. 94, pp. 362-368, 2002.

[18] Fitch, W.T., & Kramer, G. Sonifying the body electric: Superiority of an auditory over a visual display in a complex, multivariate system. In G. Kramer (Ed.), Auditory display: Sonification, audification, and auditory interfaces. (pp. 79-94). Reading, MA: Addision-Wesley. 1994.

[19] Crawley, E., Acker-Mills, B. Pastore, R. & Weil, S. Change detection in multi voice music: The role of musical structure, musical training, and task demands. Journal of Experimental Psychology: Human Performance and Perception, Vol. 28, pp. 367-378, 2002.

[20] Gregory, A. Listening to polyphonic music. Psychology of Music, Vol. 18, pp. 163-170, 1990.

[21] Sloboda, J. The musical mind. The cognitive psychology of music. Oxford UK: Clarendon Press. 1985.

[22] Neuhoff, J., Kramer, G. & Wayand, J. 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.

[23] Jones, M., Boltz, M. & Kidd, G. Controlled attending as a function of melodic and temporal context. Perception and Psychophysics, Vol. 32, pp. 211-218, 1982.

[24] Melara, R. & Marks, L.E. (1990). HARD and SOFT interacting dimensions: Differential effects of dual context on classification. Perception and Psychophysics, Vol. 47, pp. 307-325, 1990.

[25] Melara, R.D. & Marks, L.E. Interaction among auditory dimensions: timbre, pitch, and loudness. Perception and Psychophysics, Vol. 48, pp. 169-178, 1990.

[26] Melara, R.D., & Marks, L.E. Perceptual primacy of dimensions: Support for a model of dimensional interaction. Journal of Experimental Psychology: Human Perception and Performance, Vol. 16, pp. 398-414, 1990.

[27] Mondor, T.A., Zatorre, R.J., & Terrio, N.A. Constraints on the selection of auditory information. Journal of Experimental Psychology: Human Perception and Performance, Vol. 24, pp. 66-79, 1998.

[28] Bregman, A. Auditory scene analysis. Cambridge, MA: MIT Press. 1990.

Design Task	**Information required**	**Outstanding research questions**
Understand the work domain	What are the purposes and principles of the work domain? What tasks are performed? Who performs the tasks? What data are used to perform the tasks	What analysis techniques are appropriate?
Represent higher order relationships	Analysis of expert performance in the work domain What data patterns are important? What higher order relationships are important?	How can meaning be communicated effectively in sound? How can higher order relationships be represented in sound? Are there auditory emergent features that can be exploited to draw attention to significant changes? How can a sonification support both expert and novice performance?
Determine the number of variables to be displayed aurally	Analysis of the work domain and the tasks to be performed	How many variables can be monitored aurally? Decision taxonomy for determining which variables should be monitored aurally
Scale the auditory dimensions and the data variables	How perceptible is the dimension? How many steps can be represented by this dimension? How large do the increments in this dimension need to be in order to be noticeable? Does the dimension have a linear quality? What does the operator need to know about the variables? For example, is it sufficient to know that a variable is outside the normal range, or is information about how far outside normal required?	How quickly and accurately can people perceive changes in the dimensions that could potentially be used for sonification? How much change is required in a dimension to attract attention? Which dimensions can be used to create a linear order of values? How does the direction of a change in the sound dimension affect perception?
Map data to sound dimensions	Take into account the constraints of the variables and the parameters (eg. Range, linearity). Consider whether perceptual interactions are likely to occur between dimensions (could be positive or negative).	What is the potential for perceptual interactions between all dimensions used for sonification? What causes reduced or increased monitoring accuracy when multiple dimensions change? What design strategies might reduce perceptual interactions between concurrently changing dimensions? What training strategies might increase monitoring accuracy when multiple dimensions change?
Determine the number of auditory streams	Will the sonification be in focal attention or will attention be divided between the sonification and another task? What dimensions will be used to carry information? Which dimensions will interact? (could be positive or negative) Which data variables are related and should be grouped together?	How many streams can be monitored effectively? What is the effect of the dimensions on how many streams can be monitored? What effect does attentional set have on the number of streams that can be monitored? What is the most effective way to create perceptually separate auditory streams? Is monitoring aided by grouping related data variables on separate streams?
Map auditory dimensions to streams	Take into account the constraints of the variables and the sound dimensions. Will unique dimensions be mapped to each stream? Which dimensions will interact? (could be positive or negative)	Are repeated mappings of dimensions to streams more or less effective than unique mappings? How important is it to group related variables onto streams?

Table 1 Empirical framework for sonification design