DESIGNING SONIFICATION FOR EFFECTIVE ATTENTIONAL CONTROL IN COMPLEX WORK DOMAINS

Janet Anderson and Penelope Sanderson
ARC Key Centre for Human Factors and Applied Cognitive Psychology
The University of Queensland
St. Lucia, Queensland, Australia

Complex safety-critical work domains such as anesthesia require human operators to direct their attention appropriately. A sonification is a possible method for directing attention to relevant changes while still allowing monitoring under divided attention conditions. However, there is currently little information to guide the design of sonification. Two experiments investigated the effect of the number of auditory streams on ability to detect changes (Experiment 1), and the effect of the number of auditory streams under different attention conditions (Experiment 2). When monitoring with selective attention, participants noted changes more accurately with three streams than with one or two streams, but when there were also distracter changes participants noted changes more accurately in multiple streams. Overall, accuracy was lower when attention was divided than when it was selective, but accuracy was especially low in the three-stream configuration. Distracter changes increased divided attention accuracy. The results suggest that the number of streams should be minimized if operators' attention will be divided between monitoring and other tasks.

INTRODUCTION

In many safety critical environments, it is often necessary to alert the human operator to state changes. Because the visual channel is often overloaded, researchers are increasingly exploring auditory and haptic displays for this purpose (Sklar & Sarter, 1999; Watson & Sanderson, 2004).

The work reported in this paper was motivated by the search for effective continuous auditory displays—*sonification*—for anesthesia monitoring. A sonification is a continuous stream of auditory information in which changes in the variables being monitored are represented by changes in the dimensions of the sound, such as pitch. The basic literature revealed little directly relevant information to help us support the anesthesist's attention with sonification. In this paper we report two studies that continue our work at filling this gap in the basic literature (Anderson & Sanderson, 2002). The studies investigate the effect of different sonification configurations on people's ability to detect changes in auditory stimuli under selective and divided attention conditions. In these studies, *sonification configuration* refers to the number of auditory streams that are used to carry information. We manipulated the number of auditory streams in Experiment 1 and the number of auditory streams and attentional set in Experiment 2. The anesthesia domain provided the constraints within which the experimental stimuli were designed. However, the general aim of the studies was to provide data that would guide sonification design in any complex domain in which operators monitor many data streams, and timeshare multiple tasks.

Background. Sonification has been suggested as a method of presenting information to operators even when their visual attention is directed elsewhere (Woods, 1995; Seagull, Xiao, Mackenzie & Wickens, 2000). Auditory presentation of information might allow operators to monitor multiple data streams more effectively than visual presentation. The assumed rationale for this is that the auditory system is able to monitor the sound pre-attentively and draw attention to changes in the variables immediately (Woods, 1995). A continuous link is therefore maintained with the data, enabling early detection of evolving problems.

For a sonification to be effective, information should be able to be easily interpreted under both divided and selective attention conditions. In a complex environment, operators are required to multi-task and must allocate their attention effectively to ensure they don't miss unexpected changes. For example, in a study of the monitoring tasks carried out by anesthetists, Seagull et al. (2000) found that anesthetists divide their attention across all information displays to detect any change in the patient's physiological state, and also selectively attend to particular variables that they need to assess and track. The challenge is therefore to design a sonification that attracts attention to changes but also supports selective attention.

Unfortunately, the basic information needed to design a sonification to control attention appropriately is not available in the auditory attention literature (Sanderson, Anderson & Watson, 2000) and is only starting to emerge in the sonification literature (Kramer, 1994). Many studies in the attention literature use simple, static stimuli that are presented in discrete trials. The results cannot be generalized to stimuli that are complex and change over time (Neuhoff, McBeath & Wanzie, 1999). Auditory Scene Analysis research, however, does suggest some potential problems (Bregman, 1990). Research conducted in this paradigm investigated the conditions under which an auditory streaming illusion occurs. The illusion occurs when a single sound that fluctuates in pitch appears to separate into a high pitched stream and a low-pitched stream. The research shows that auditory streams form coherent perceptual objects that can be attended and followed over time. When there are multiple streams present, only one stream can be in focal attention and attention must be switched between streams (Handel, 1989; Bregman, 1990). This raises the possibility that in multiple stream sonifications information in one stream might be missed if attention is focused on a different stream (Kramer, 1994). If so, a single stream sonification might be more effective than multiple streams. The present studies were motivated by the need to provide empirical information that would be useful to sonification designers.

Sonifying physiological variables. The auditory display of physiological variables in anesthesia has a long history. Pulse oximetry, an auditory display of heart rate and arterial oxygen saturation, is routinely used in anesthesia for patient monitoring. This display is a continuous series of beeps with heart rate encoded to the speed of the beeps and oxygen saturation encoded to the pitch of the beeps. Recently, human factors researchers in anesthesia have reported on the development of more complex sonifications that include additional physiological variables, thus extending the existing pulse oximetry sonification. Fitch and Kramer (1994) developed a two-stream sonification of eight physiological variables. They trained participants to identify an abnormal physiological event and compared monitoring accuracy for the sonification with a visual display and a combined auditory and visual display. Best performance was attained with the sonification. In later work, either five or six physiological variables were sonified using two auditory streams (Loeb & Fitch, 2002; Watson & Sanderson, 2004). In their six-variable sonification, Loeb and Fitch (2002) mapped pitch, speed and timbre to each of two auditory streams. Watson and Sanderson (2004) developed a five variable display and mapped pitch, speed and loudness to one stream and pitch and speed to the other stream. The theoretical rationale for favoring a two-stream configuration with repeated mappings of dimensions to streams was not discussed in the above literature.

The role of attention in monitoring a sonification has also been explicitly addressed in previous studies. Several studies have found that when participants were dividing their attention between performing another task and monitoring physiological data, they performed the other task more accurately when the physiological data were presented auditorily compared with visually (Watson and Sanderson, 2004; Sanderson et al., 2004; Seagull et al., 2000). Clinical anesthetists monitored the patient as accurately with sonification as with visual displays. Overall, sonification helped participants time-share their tasks, suggesting that sonification allows effective monitoring of multiple data streams.

Sonification stream configuration. However, the question of how the configuration of the sonification affects people's ability to monitor a sonification under both selective and divided attention appears not to have been addressed. In particular, the effect of the number of auditory streams has not been investigated. Although previous research has favored a two-stream configuration, it is theoretically possible to configure a sonification with any number of streams. For a six-variable display, the number of streams could range from, for example, a one-stream configuration with six changing dimensions to a six-stream configuration with one dimension changing on each stream. A three-stream configuration probably represents a practical limit. In a sonification of six physiological variables, the six variables could be mapped to one stream with six changing dimensions, two streams with three dimensions changing on each stream or three streams, with two dimensions changing on each stream.

It is currently unclear how sonification designers might assess which is the best configuration. The configuration of a sonification might interact with attention. Visual display formats that support divided attention tasks make it difficult to selectively attend to a particular part of the display (Wickens & Hollands, 2000). These results have been found with Chernoff face displays (Suzuki & Cavanagh, 1995) and object displays, in which dimensions of the objects such as height and shape are used to represent data variables (Carswell & Wickens, 1987; Mori & Hayashi, 1995). It is unclear to what extent visual design principles can be applied to auditory displays.

	Stream Configuration					
	One Stream	**Two Streams**		**Three Streams**		
Dimension	**Stream A**	**Stream A**	**Stream B**	**Stream A**	**Stream B**	**Stream C**
Amplitude	Changing	Constant	Changing	Constant	Constant	Changing
Frequency	Changing	Changing	Constant	Changing	Constant	Constant
Harmonics	Changing	Constant	Changing	Constant	Changing	Constant
PulseSpeed	Changing	Changing	Constant	Changing	Constant	Constant
Tremolo	Changing	Changing	Constant	Constant	Changing	Constant
PulseWidth	Changing	Constant	Changing	Constant	Constant	Changing

Table 1. Design of stream configurations showing which acoustic dimensions change on each stream.

EXPERIMENT 1: STREAM CONFIGURATION

The main question addressed in Experiment 1 was whether the number of auditory streams affected people's ability to selectively attend to acoustic dimensions and identify when they changed. It is not clear from the auditory attention literature how many simultaneous auditory streams can be attended. The question was partly addressed by Brochard, Drake, Botte and McAdams (1999) who used the auditory scene analysis paradigm to investigate listeners' ability to selectively attend to one stream embedded in a complex auditory stimulus of up to three other streams. They found that attending to one stream when it was embedded in a number of other streams was more difficult than attending to a single stream. This suggests that a sonification with multiple streams might be less effectively monitored than a single stream.

The aim of Experiment 1 was to investigate the effect of stream configuration on accuracy in a change detection task. The three stream configurations were:

- One stream with six changing dimensions
- Two streams with three changing dimensions
- Three streams, each with two changing dimensions.

The same acoustic dimensions were used for the one-stream, two-stream and three-stream configurations so that any differences between the sonification configurations could be attributed to the different numbers of streams rather than to differences in the discriminability of the dimensions used.

Method

A pulsing sound was used for each auditory stream, with six changing auditory dimensions: amplitude, frequency, harmonics (selected by a band pass filter), pulse speed, tremolo (amplitude modulation within the pulse) and pulse width. Dimensions were uniquely mapped to each stream, as shown in Table 1.

All twelve participants were university students who received course credit. They monitored 10-second long sound sequences for a change in the target dimension. In the no distraction condition only the target dimension changed. Participants were instructed to selectively attend to the target dimension. In the one distracter condition one other dimension also changed at the same time as the target dimension changed, creating a more difficult selective attention task.

Results

Both stream configuration and the presence of distractors exerted a significant effect on performance accuracy. A three-way within subjects ANOVA found that the main effects for stream configuration, $F(2,22)=7.144$, $MSE=.017$, $p=.004$, distractor $F(1,11)=9.38$, $MSE=.013$, $p=.010$, and dimension $F(5,55)=.34.89$, $MSE=.024$, $p<.00001$, were all significant. There was also a significant three-way interaction between stream configuration and distractor $F(2,22)=7.85$, $MSE=.010$, $p=.002$ and between stream configuration and dimension $F(10,110)=4.43$, $MSE=.012$, $p=.000031$. Accuracy was significantly lower for the three-stream configuration ($M=.71$), compared to the one stream ($M=.76$) and two stream ($M=.76$) configurations. Distraction significantly reduced accuracy for the one stream configuration, $F(1,11)=28.97$, $p=.0002$, but not for the two stream $F(1,11)=3.10$, $p=.11$ or three stream configurations, $F(1,11)=.14$, $p=.71$.

In summary, participants' accuracy at detecting changes in sonification was influenced by the number of streams in the configuration, with the single-stream sonification supporting best performance when one dimension at a time changed. If there was a distractor change at the same time, accuracy was not different for the different stream configurations.

EXPERIMENT 2: STREAM CONFIGURATION AND ATTENTION

The aim of Experiment 2 was to test whether monitoring accuracy for the three different stream configurations would be comparable under selective and divided attention conditions. This is an important question because the hypothesized benefit of data sonification is that attention would be attracted by changes in the sound (Woods, 1995). Previous studies have shown that sonification is effective when attention is divided (Watson & Sanderson, 2004) but no studies have investigated the joint effects of sonification configuration and attentional set. A study using musical stimuli suggests that different stimulus configurations might support selective and divided attention. Crawley, Acker-Mills, Pastore & Weil (2002) investigated change detection in three-voice musical melodies under selective and divided attention. They found an interaction between attention and the type of stimulus (simultaneous or asynchronous note onsets). Some types of sequences facilitated divided attention by making the change very noticeable, but the very feature that supported divided attention made selective attention more difficult. However, they did not investigate the effect of different numbers of musical voices on change detection accuracy.

In Experiment 2 change detection accuracy was assessed for each stream configuration when it was in selective attention and when attention was divided between auditory monitoring and a secondary task.

Method

Experiment 2 used the same stream configurations as those in Experiment 1. The stimuli were 20-second long sequences of pulsing sound. Changes were heard either with no distraction or with one distracter change. Twelve participants monitored the sounds in either selective attention or divided attention. In the selective attention condition, trials were blocked according to the target dimension and participants were cued to attend to a particular dimension. In the divided attention condition they simultaneously monitored the sound and performed a forced pace arithmetic task presented on another monitor.

Results

Accuracy was analyzed in a four-way repeated measure ANOVA. There were significant main effects for stream configuration $F(2,22)=12.15$, $p=.00002$ and dimension F(5,55=14.14, $p<.00001$, but not for distracter or attention. The interaction between attention and stream configuration and all the three way interactions that included stream configuration and attention were non-significant.

In contrast, the interaction between attention and distractors was highly significant, $F(1,11)=62.84$, $p<.00001$. Accuracy was improved by the presence of distractor changes when attention was divided, but reduced by distractor changes when participants were selectively attending to a target dimension.

The four-way interaction was further probed using analysis of simple effects. When the effect of stream configuration was examined at each level of attention, in the divided attention task, stream configuration had the effect of reducing accuracy as the number of streams increased and this was consistent across distractors and dimensions. For selective attention, the effect of stream configuration on accuracy occurred through interactions with distractors and dimensions.

DISCUSSION

Stream configuration. Taken together, the results of the two experiments showed that monitoring a single auditory stream was more accurate than monitoring multiple auditory streams when only one dimension changed. When two dimensions changed simultaneously, however, monitoring multiple streams was as accurate as monitoring a single stream because interacting dimensions can be mapped to different streams, reducing the detrimental effects of perceptual interactions.

The lower accuracy for the three-stream configuration in Experiment 1 can be attributed to lower accuracy for the harmonics, speed and tremolo dimensions. In contrast, accuracy was actually higher for amplitude in the two and three stream configurations than in the one-stream configuration. As for Crawley et al.'s (2002) finding that dissonance between streams facilitated change detection performance, this was probably because the difference between the amplitudes of the streams became noticeable as the amplitude of one stream changed, facilitating selective attention.

Attention. Divided attention monitoring was less accurate than selective attention monitoring when there was no distraction, and monitoring three auditory streams in divided attention was especially difficult. When attention was divided, changes in multiple stream configurations were not as attention attracting as changes in a single stream. Multiple streams apparently made it harder for a change to pop out of the noise and attract attention. The least discriminable dimensions—tremolo, width, and to a lesser extent speed—were particularly affected

Simultaneous changes in two dimensions had the effect of making changes much more noticeable when attention was divided, as would be expected from research showing that processing combinations of auditory dimensions can be more efficient than processing single dimensions (Boltz & Jones, 1986; Mondor, Zatorre & Terrio, 1998). Participants in this experiment were not asked to identify what had changed in the divided attention task so it is not known whether the increased accuracy in the one-distractor condition was because they were responding only to the most discriminable dimension or to the overall effect of simultaneous changes in more than one dimension.

Conclusion. The purpose of these experiments was to provide some empirical data to guide the process of sonification design, with a particular concern for supporting both selective and divided attention. To ground the study, some properties of anesthesia monitoring were captured—specifically, the number of variables, their numerical ranges, and the kind of discriminations required (Watson & Sanderson, 2004). The results showed that in the presence of distraction, multiple streams provide support for selective attention by reducing perceptual interactions. However, when attention is divided, multiple streams are very difficult to monitor, especially when only one dimension changes. This suggests that designers should reduce the number of auditory streams if sonification will be monitored outside selective attention.

In these studies, we investigated the effect of perceptual processes on people's ability to detect changes in auditory stimuli. Our motivation was to fill a gap in the basic literature, and draw implications for design. Further factors are crucial, though, and our job is far from done. For example, domain experts will form expectations about system state based on the past and present structure and interrelationships of the data, the expert's knowledge, and what the expert is trying to achieve. Such expectations will no doubt introduce further factors to the picture we have presented here. It is therefore important to examine the interplay of perceptual processes and strategic processes when attention is guided by sonification. Accordingly, future experiments will employ domain experts and data streams with interrelations closer to those seen in anesthesia.

REFERENCES

Anderson, J., & Sanderson, P. (2002). Sonification design for real-time processes: Issues and a demonstration. *Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society*. Baltimore, MD. September 30-October 4.

Boltz, M. & Jones, M. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology, 18,* 389-431

Bregman, A. (1990). *Auditory scene analysis.* Cambridge, MA: MIT Press.

Brochard, R., Drake, C., Botte, M.C. & McAdams, S. (1999). Perceptual organization of complex auditory sequences: Effect of number of simultaneous subsequences and frequency separation. *Journal of Experimental Psychology: Human Performance and Perception, 25,* 1742-1759.

Carswell, C.M. & Wickens, C.D. (1987). Information integration and the object display: An interaction of task demands and display superiority. *Ergonomics, 30,* 511-527.

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

Fitch, W. T., & Kramer, G. (1994). 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: Addison-Wesley.

Kramer, G. (1994). 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.

Loeb, R.G., & Fitch, W.T. (2002). A laboratory evaluation of an auditory display designed to enhance intra-operative monitoring. *Anesthesia and Analgesia, 94,* 362-368.

Melara, R.D., & Marks, L.E. (1990b). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 398-414.

Mondor, T.A., Zatorre, R.J., & Terrio, N.A. (1998). Constraints on the selection of auditory information. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 66-79.

Mori, H. & Hayashi, Y. (1995). Visual interference with users' tasks on multiwindow systems. *International Journal of Human-Computer Interaction, 7,* 329-340.

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

Sanderson, P., Anderson, J., & Watson, M. (2000). 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)*. CSIRO: Sydney. Pp 259-266.

Sanderson, P., Crawford, J., Savill, A., Watson, M., & Russell, W. (2004). Visual and auditory attention in patient monitoring: A formative analysis. *Cognition, Technology and Work , 62(3)*

Seagull, F.J., Xiao, Y., Mackenzie, C., & Wickens, C. (2000). Auditory alarms: From alerting to informing. *Proceedings of the joint meeting of the Human Factors and Ergonomics Society and the International Ergonomics Association (IEA2000/HFES2000).* Santa Monica, CA: HFES, 1, 223-226.

Sklar, A., & Sarter, N. (1999). Good vibrations: Tactile feedback in support of attention allocation and human-automation coordination in event-driven domains. *Human Factors*, 41(4), 543-552.

Suzuki, S. & Cavanagh, P.. (1995). Facial organization blocks access to low-level features: An object inferiority effect. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 901-913.

Watson, M. & Sanderson, P. (2004). Sonification helps eyes-free respiratory monitoring and task timesharing. *Human Factors,* 46(3).

Wickens, C. & Hollands, J. (2000). *Engineering psychology and human performance*. New York: Harper Collins.

Woods, D.D. (1995). The alarm problem and directed attention in dynamic fault management. *Ergonomics, 38,* 2371-23