RICE UNIVERSITY

Dimensions of Sound in Auditory Displays: The Effects of Redundant Dimensions

by

S. Camille Peres

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Master of Arts

APPROVED, THESIS COMMITTEE:

____
David M. Lane, Associate Professor, Chair
Psychology and Statistics

____
James R. Pomerantz, Professor
Psychology

____
Michael D. Byrne, Assistant Professor
Psychology

Houston, Texas

AUGUST 2004

Abstract

Dimensions of Sound in Auditory Displays:

The Effects of Redundant Dimensions

by

S. Camille Peres

Three experiments are presented comparing the effectiveness of several parameters of sound for the auditory presentation of statistical data. The dimensions of pitch, loudness, panning, and time were used alone and redundantly to map the values of a box plot to an auditory display. Temporal mappings resulted in better performance than mappings using pitch, panning, or loudness. In the first two experiments, there was no benefit when the mapping condition used two dimensions redundantly over mappings using one dimension. However, for the third experiment, there was a benefit of a redundant design when the dimensions of sound used were integral whereas there was no benefit when they were separable. This third experiment used a task more closely approximating a real-life application of auditory displays. Its results suggest that sonification can be used effectively in situations requiring the monitoring of more than one source of information.

Acknowledgments

I would first like to thank my advisor, David Lane for guiding me through this process. He always offered guidance and insight in the most encouraging and patient manner, never holding me back, but never letting me get too far astray. Others have attempted but with much less success. The wonderful conversations we have about research, politics and the virtues of box plots are some of my favorites. Additionally, I would like to thank the other members of my committee, Jim Pomerantz and Michael Byrne, for their critical feedback and support. It has been a privilege working with all of these people.

More gratitude goes to my colleague and friend, Anikó Sándor. The conversations we've had about sonification, graduate school, and life have been wonderful. Others instrumental to the success of this project were: Joan Lu, for the Java programming; Jeremyah Payne, for help with creating the sounds; and Adi Bichman, for help with data collection and analysis.

And finally, I would like to recognize my extraordinary support network. My efforts would not have been possible with out their constant and continuing support. Their acts of kindness and support are too numerous to list, but I do want to mention the following people: my friends and neighbors; Dad, Linda and Kendra, who always listen to me talk about my research no matter how obscure it is; Mom and Herman, whose help and love was pivotal in my completion of this research; and my children, Landon and Audrey, who continue to inspire me everyday. Finally, my most heartfelt gratitude goes to my husband, who always manages to convince me that I can do it, even when I have convinced myself I can't. He continues to big my biggest fan.

Table of Contents

List of Tables

Dimensions of Sound in Auditory Displays:
The Effects of Redundant Dimensions

Statistical graphs are important for communicating parameters of a dataset. It is well documented that normal visual presentation of graphs can be very effective (Cleveland, 1994; Cleveland & McGill, 1986). However, a visual presentation is not practical on devices without displays or with small displays (i.e., cell phones and PDAs), for individuals with visual disabilities (Petrie & Morley, 1998; Walker & Lane, 2001), and for divided attention tasks when the user's eyes are otherwise occupied (Kramer, 1994). Sonification, or the representation of data through sound or non-speech audio, could be effective in these situations. This paper focuses on the use of sonification for one type of statistical graph: box plots. Box plots are widely used and important graphical displays. Moreover, their simplicity makes them well suited for testing the basic principles of sonification being investigated.

John Flowers and his colleagues conducted a series of studies comparing the perception of auditory and visual presentations of statistical graphs (Flowers, Buhman, & Turnage, 1997; Flowers & Hauer, 1992, 1993, 1995). They used different dimensions of sound both independently and together to represent the parameters of a data set. For instance, a frequency polygon was represented by using pitch to represent the Y-axis and loudness for the values on the X-axis. Of interest was the relationship between statistical properties of data and similarity judgments of visual and auditory graphs. In general, they found that the perception of auditory and visual displays was similar: judgments of both types of graphs were influenced by skew, spread, and central tendency. However, judgments of visual graphs were relatively more influenced by skew whereas judgments of auditory graphs were relatively more influenced by central tendency.

This pioneering research by Flowers et al. provides pertinent information regarding auditory displays of statistical data. However, many issues regarding the designs of these displays remain unresolved. One issue involves the relative effectiveness of various sound dimensions for representing data. Many of auditory displays used by Flowers et al. utilized the dimension of pitch, but the performance with these auditory graphs was never compared to performance with graphs utilizing other dimensions. Furthermore, although research suggests that spatial location and temporal designs may be effective dimensions for use in auditory designs (Lorho, Marila, & Hiipakka, 2001; Watkins, LeCompte, Elliott, & Fish, 1992), there is relatively little data on this topic. One of the goals of the first two experiments was to compare the relative effectiveness of displays using the temporal aspects of sound, spatial location (panning), loudness, and pitch. Another goal of these experiments was to investigate the effects of using dimensions redundantly in auditory displays. There is reason to believe that in some contexts, using two sound dimensions together in a redundant fashion is a better representation of the data than could be achieved by the dimensions used individually (Kramer, 1996). The third experiment was designed to continue the investigation on redundant dimensions and specifically compared auditory displays using integral and separable dimensions of sound. In speeded card sorting tasks, integral dimensions of sound result in better performance when two dimensions are used redundantly and Experiment 3 was designed to see if this would generalize to a divided-attention monitoring task. Finally, I was interested in the subjective impressions of auditory graphs. It is often assumed that participants will prefer conditions where they have a better performance, but previous research has provided counter examples (Petrie & Morley, 1998).

EXPERIMENT 1

This experiment compared the effectiveness of different dimensions of sound for presenting the statistical information normally contained in box plots. Specifically, the box plot represented the "five number summary" of a distribution, which consists of the minimum, 25^{th} percentile, median, 75^{th} percentile and the maximum of the distribution. Pitch, panning, and a redundant design with both pitch and panning were the sound dimensions compared in this experiment.

Method

Task. Participants listened to a sonified box plot and then selected a visual representation of the box plot from a set of four. The box plot they selected would play and participants were told whether or not their choice was correct. If they were incorrect, they continued selecting box plots until they had selected the correct one. For purposes of data analysis, the response on a trial was considered correct only if the participant selected the correct graph on the first try. If the participant selected the incorrect box plot on the first try, it was considered an error. Participants were instructed to listen to the auditory box plots as many times as they wanted to before selecting a visual box plot; the response times for the trials were not recorded.

Stimuli. The stimuli consisted of twenty box plots that varied in skew, location (central tendency), and spread. On a given trial, one of the three distractors differed from the target in skewness, one in either location or spread (determined randomly), and if this distracter differed in location, then the third distracter differed in both skew and location. If the second distracter differed in spread, then the third distracter differed in skew and spread. Within this framework, the target and the distracters were determined randomly on each trial.

There were three different sound dimension conditions: pitch, panning, and a redundant condition (pitch and panning). For the pitch condition, the values from the box plots were mapped to a note in the range of 16 notes below and 16 notes above 440 Hz (middle C). For the panning condition, the values from the box plots were mapped to an amplitude scale that lateralized the values to points in space on an axis that goes through the ears. The redundant condition used both the pitch and panning transformations. The sounds were played in the following sequence: the absolute minimum of the scale, the values of the box plot (minimum, lower 25th, median, upper 75th, and maximum), and the absolute maximum of the scale. The box plot sounds were played without pauses and there was a 1-second pause separating the absolute minimum and maximum values from the other sounds.

Participants. Fifty-six undergraduate students between the ages of 18 and 24 participated in the experiment for course credit. There were 47 females and 9 males and each were randomly assigned to the three conditions with the constraint that the groups were as equal in size as possible given the total sample size. The pitch and redundant conditions each had 19 participants; the panning condition had 18 participants.

Procedure. After a training session, each participant completed 50 trials. The training session took approximately fifteen minutes and the experiment took an average of 28 minutes (range: 12.7 to 50.4 minutes). After the experiment, the participants provided some demographic information and answered several questions regarding their impression of the auditory box plots and the task. Participants used a five-point scale to rate the sounds on how pleasant or annoying there were and used a similar five-point scale to rate how difficult the task was. For all three scales, 1 was the most positive rating and 5 the least positive.

Macintosh computers with Internet access in a laboratory environment were used for the experiment and an interactive web site programmed with JavaScript was used for training and data collection. Sounds were presented to participants over headphones.

Results

The trials were grouped into five blocks of ten and the mean proportion correct for the blocks are shown in Table 1. This table shows that the participants improved at the task and that the learning occurred primarily in the first block of ten trials, reaching an asymptote of only 0.50 (with chance being 0.25).

Table 1
Mean proportion correct for the five blocks of trials.

	Block 1	Block 2	Block 3	Block 4	Block 5
Mean Proportion Correct	0.44	0.50	0.49	0.50	0.53

The linear component of trend was significant, $F\,(1,55) = 5.59$, $p = 0.025$. Note that a significant linear component does not mean that the relationship is linear. There was no evidence that the shape of the learning curve differed as a function of Sound Dimension, $F\,(2,53) = 0.21$, $p = 0.815$.

The distributions of the proportion correct as a function of sound dimension are shown in Figure 1. The pitch condition had the highest average proportion correct

(M=0.58) and the panning condition had the lowest (M=0.37). The redundant condition (M=0.50) was intermediate. The differences among the conditions were significant, *F* (2,53) = 10.82 $p < .001$, and the Tukey HSD ($p < 0.05$) showed that the panning condition was significantly lower than the other two conditions which did not differ significantly from each other.

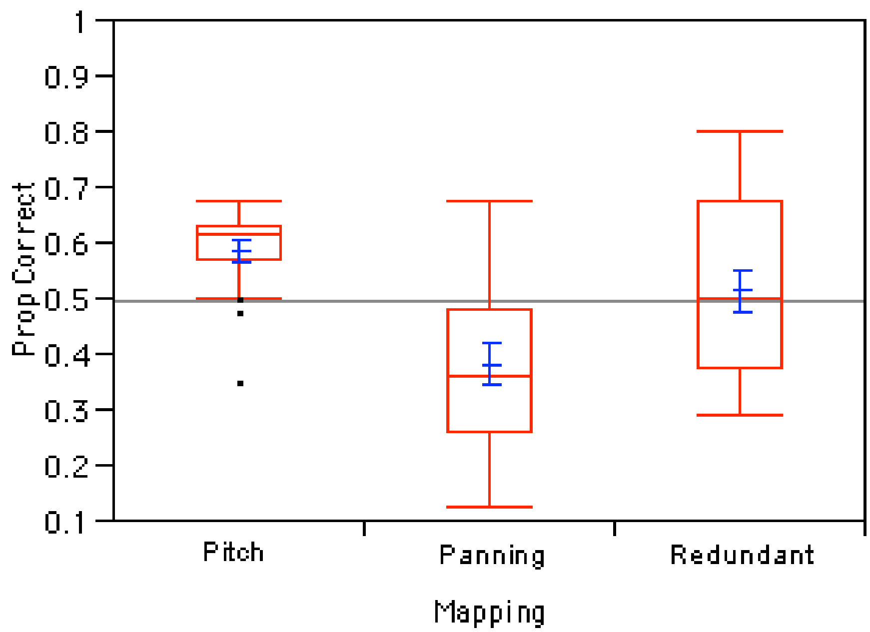

Figure 1 Box plots of proportion correct as a function of Sound Dimension. In addition to the median, each box plot shows the mean ± 1 standard error. The horizontal line traversing the graph indicates the grand means.

Participants' choices on error trials revealed the aspect of the distribution on which the error was based. For example, if a participant chose a box plot that was identical to the target in all respects except for location, this was deemed a location error. Analysis of the errors (adjusting for the fact that there were more opportunities to make skew errors than other errors) showed a lower proportion of location errors than either spread or skew errors. Table 2 provides the means as a function of condition and shows that the proportion of spread errors did not differ as a function of sound dimensions,

$F(2,52) = 1.09$, $p = 0.343$. However there were more errors for panning condition for both skew, $F(2,52) = 18.48$, $p < 0.001$, and location $F(2,52) = 7.52$, $p < 0.01$. A Tukey HSD test showed significant differences ($p < 0.05$) between the panning condition and each of the other two that did not differ significantly from each other.

Table 2
Mean proportion of the three errors types as a function of Sound Dimension.

	Pitch	Panning	Redundant
Spread	0.30	0.35	0.32
Skew	0.26	0.46	0.34
Location	0.13	0.30	0.20

As can be seen in Table 3, participants rated the redundant condition as less difficult, less annoying, and less unpleasant than the other conditions ($p < 0.001$ for all comparisons). Note that although the participants had a strong preference for the redundant condition, performance did not mirror this preference.

Table 3
Experiment 1: Mean subjective ratings (with 1 being the most positive rating and 5 being the least).

	Pitch	Panning	Redundant
Difficulty	2.21	2.94	1.95
Annoying	2.42	3.06	1.05
Unpleasant	2.16	2.33	1.00

Section 1.01

Summary

Participants found the task difficult, and while they improved over the first 10 trials, there was no subsequent improvement. The mean proportion correct was about

0.50 (chance was 0.25), with considerable variability in the participants' performances: the proportion correct ranged from 0.13 to 0.90. The pitch condition had the best performance, although it was not significantly better than the redundant condition. Performance in the panning condition was worse than in either of the other conditions.

There was an interesting disassociation between subjective impressions and performance. Subjective impressions in the redundant condition were considerably higher than for the pitch condition even though performance in the redundant condition was slightly (though not significantly) lower than in the redundant condition.

EXPERIMENT 2

This first experiment provided interesting results on the effects of different sound dimensions when used for sonifying box plots. However, there remained questions regarding whether the effects of a redundant design would change when different sound dimensions were used. Furthermore, there is some evidence that the temporal dimension of sound could be beneficial in auditory displays such as the ones used for Experiment 1. In 1992, Watkins and his colleagues found that memory for an irregular temporal pattern was better with auditory stimuli than with visual stimuli (Watkins et al., 1992). This suggests that the use of the temporal dimension of sound may facilitate the identification of an auditory box plot.

Thus, the second experiment was conducted to extend the findings of Experiment 1 and compare the effectiveness of a temporal design with that of a pitch, and redundant design using both pitch and temporal dimensions of sound for representing the "five-number-summary" box plots. Furthermore, in Experiment 1 performance appeared to asymptote very quickly. In experiment 2, participants completed 100 rather than 50 trials to determine if they had actually reached an asymptote, or whether further practice would lead to further increases in performance. Finally, in Experiment 2, the presentation time used to play a box plot was manipulated, one time being twice as long as the other. The box plots in Flower's et al. study were approximately half the length of the box plots in Experiment 1. It is conceivable that the shorter box plots would be perceived as a unit rather than 7 separate sounds. Thus, the length of the box plot was manipulated to investigate any effects the different lengths may have on performance.

Method

Design. A Sound Dimension (pitch, temporal, and redundant) x Presentation Time (short and long) factorial design was employed. Both Sound Dimension and Presentation Time were manipulated as between-subjects variables.

Task. Participants were presented with a task identical to that in Experiment 1. The pitch condition was mapped using the same method as in Experiment 1. In the temporal condition, the distances between the values of the box plot were represented by the time between the onsets of the sounds and the pitch of all the sounds remained constant at 440 Hz. For the redundant conditions, the values of the box plots were mapped using both the pitch and the temporal mappings. The sounds of the box plots were played in the same sequence as Experiment 1. The box plots with the Long-Presentation-Time played for 9 sec. while those with the Short-Presentation-Time played for 4.5 sec. This resulted in six conditions: pitch-long (PL), pitch-short (PS), temporal-long (TL), temporal-short (TS), redundant-long (RS), and redundant-short (RS).

Participants. Data collection occurred in two phases. The pitch and temporal mappings with the short and long presentation times were collected in the first phase (N=

38) while those conditions using the redundant mapping with the short and long presentation times were collected during the second phase (N=22). Seventy undergraduate students between the ages of 18 and 24 participated in the experiment for course credit. They were randomly assigned to one of the conditions with the constraint that the groups were as equal in size as possible given the total sample size. The data from ten participants had to be eliminated because of a problem with the interactive web site. This left a total of 60 participants, with 10 participants in the PL, TS, and RS conditions, 9 participants in the PS and TL conditions, and 12 participants in the RL condition. There were 32 females and 27 males and one participant did not indicate his/her gender.

Procedure. Participants completed 100 trials that took approximately one hour. The testing design and environment were identical to Experiment 1 and, as in the first experiment, participants answered several questions regarding their impressions of the box plots along with demographic information.

Results

The trials were grouped into 10 blocks of ten and the proportion of correct responses increased from 0.53 to 0.61 over the 10 blocks. The linear component of trend was significant, $F(9,46) = 2.41$, $p = 0.024$ and performance appeared to reach asymptote at the fifth block. The linear component of trend did not interact significantly with either Sound Dimension, $F(18,92) = 0.90$, $p = 0.579$, or Presentation Time, $F(9,46) = 1.89$, $p = 0.078$. Nor was there a Trials (linear) x Sound Dimension x Presentation Time interaction $F(18, 92) = 1.54$, $p = 0.094$.

Skew errors decreased significantly over the 100 trials whereas the proportion of spread and location errors did not. For skew and location errors, the shape of the learning curve was different for the two presentation times. The proportion of skew errors decreased from 0.38 to 0.24 in the short condition whereas it decreased only slightly, from 0.25 to 0.22, in the long condition. For location errors, errors increased somewhat for the long condition (from 0.07 to 0.08) while they decreased for the short condition (0.13 to 0.07). Both of these Presentation Time x Trials (linear) interactions were significant; for skew errors, $F(9,46) = 2.26$, $p = 0.034$, and for location errors, $F(9,46) = 2.70$, $p = 0.013$. There were no Sound Dimension x Trial interactions for skew, location, or spread errors.

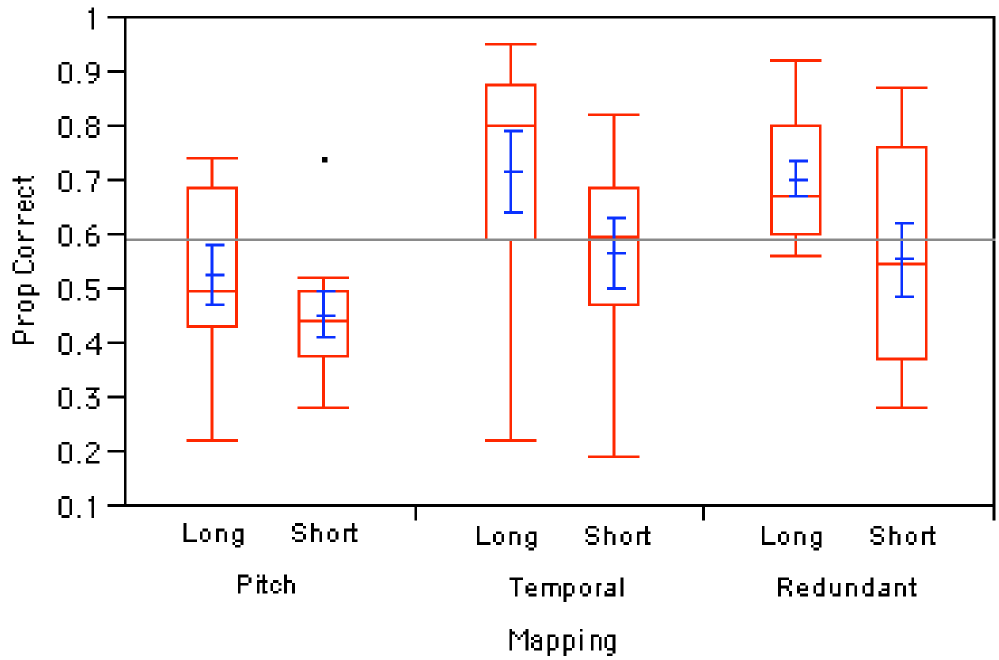

Figure 2 Box plots of proportion correct as a function of Sound Dimension and Presentation Time. In addition to the median, each box plot shows the mean ± 1 standard error. The horizontal line traversing the graph indicates the grand means.

Figure 2 provides the mean proportion correct as a function of Sound Dimension and Presentation Time and shows that performance was lower for the pitch conditions than either the temporal or redundant conditions, $F(2,57) = 3.79$, $p = 0.029$. A Tukey's pair-wise comparison showed that the pitch and the redundant conditions were significantly different while the temporal conditions were not significantly different from either the pitch or redundant conditions. Also seen in Figure 1 is the difference in performance due to Presentation Time with Long-Presentation-Time conditions having better performance than the Short-Presentation-Time conditions, $F(1,57) = 6.94$, $p = 0.011$. There was no Sound Design x Presentation Time interaction, $F(2,27) = 0.28$, $p = 0.751$.

The mean proportions of errors as a function of Sound Dimension and Presentation Time are shown in Table 4. As in Experiment 1, participants made substantially fewer location errors (M=0.15) than either spread (M=0.30) or skew errors (M=0.26). Separate Presentation Time x Sound Dimension ANOVAs were run for skew, spread and location errors. Table 4 shows that the there was little difference between the sound dimension conditions for location and spread errors. For the skew errors, however, performance was better in the temporal and redundant conditions than the pitch condition, $F(2,54) = 6.50$, $p = 0.003$. Performance in the Long-Presentation-Time condition was better than in the Short-Presentation-Time condition, $F(1,54) = 11.60$, $p = 0.001$ and there was no Sound Condition x Presentation Time interaction, $F < 1$.

Table 4
Mean proportion of the three errors types as a function of Sound Dimension and Presentation Time.

		Pitch	Temporal	Redundant
Short	Spread	0.17	0.15	0.16
	Skew	0.40	0.29	0.27
	Location	0.10	0.09	0.10
Long	Spread	0.18	0.12	0.13
	Skew	0.29	0.17	0.18
	Location	0.09	0.05	0.03

The participants rated the box plots on the same questions used in Experiment 1. As can be seen in Table 5, participants gave more favorable ratings to the temporal and redundant conditions than to the pitch condition on all three questions. However, the only significant difference occurred on the ratings of difficulty for which the conditions temporal and redundant were rated easier than the pitch dimension, $F(2,57) = 4.20$, $p = 0.02$.

Table 5
Experiment 2: Mean subjective ratings
(with 1 being the most positive rating and 5 being the least).

		Pitch	Temporal	Redundant
Difficulty	Long	2.0	0.9	1.3
	Short	2.1	1.5	1.6
Unpleasant	Long	2.5	2.0	2.1
	Short	2.4	1.8	2.3
Annoy	Long	2.6	2.1	2.0
	Short	2.3	1.8	2.5

Summary

As in the first experiment, participants found the task difficult and most reached an asymptote around the 50th trial of about 0.60 correct. Participants in the temporal and redundant conditions performed substantially higher than those in the pitch condition and this finding is consistent with the pattern of responses found in the subjective ratings. Additionally, there was no benefit of the redundant condition over the better of the other two single dimensions.

As in Experiment 1, there were fewer location errors than skew or spread errors. Also consistent with Experiment 1 was the finding that sound dimension had a larger effect on skew errors than on spread errors.

Discussion of Experiments 1 and 2

Participants generally found this task difficult and they did not get much better with practice. The best mean performance achieved in either of the two experiments was in the temporal-long condition with a mean proportion correct of 0.70. There was considerable variability in performance for both experiments with some participants performing approximately at chance and some performing *very* well. In the first experiment, the best three participants were correct 78%, 80%, and 90% of the time while the worst performances were 12.5%, 20% and 22.5%; for the second experiment, the four top performing participants were correct 87%, 87%, 88%, and 95% of the time and the worst performances were 19%, 22% and 23%. Designers of auditory statistical displays should be cognizant of this high variability.

Although Lorho et al. (2001) found that participants were relatively good at locating sounds spatially, the present data suggest that spatial location is not a good mapping technique for sonifying statistical graphs. Of the three methods examined here, temporal mapping was clearly the best. These results considered in conjunction with previous work (Watkins et al., 1992) suggest that temporal mapping can be effective for these types of displays.

Both experiments found that the proportion of location errors was substantially lower than the proportion of skew or spread errors. This mirrors the findings by Flowers and his colleagues and supports the idea that auditory displays may be especially effective for presenting information about the central tendency of a data distribution. Further information from the error analysis suggest that particular sound dimensions may be better for extracting some of the statistical parameters than others, however the measures used were somewhat indirect and would require further investigation before conclusions could be drawn. Additionally, both of these experiments used fixed effect designs and thus the findings may have limited generalizability.

For both Experiment 1 and 2, participants' preferences did not match their performance. In the first experiment, participants strongly preferred the redundant condition to the pitch condition even though performance was slightly better in the pitch condition. While in the second experiment, there were no differences in the preference ratings for the three different conditions although performance was much worse in the pitch condition than either the temporal or redundant conditions. These findings are similar to those of Petrie and Morley (1998), who also found a disassociation between performance and preference on a task using an auditory display.

Experiment 1 found that the redundant use of the panning and pitch dimensions did not result in better performance than the pitch dimension used alone; performance in the redundant condition was slightly but not significantly worse than in the pitch condition. Experiment 2 also found no benefit of a redundant design with the dimensions of time and pitch. Consistent results have been obtained (Sándor & Lane, 2003) using a simple task in which participants map sounds to absolute numeric values. This research found that performance was significantly lower when temporal and pitch information were presented redundantly than when temporal information was presented alone. This provides further evidence that, for the dimensions used here, redundancy is not better and may be worse than the better of the two individual dimensions in this type of task. Naturally, there may be other contexts and/or other dimensions for which an auditory design using dimensions of sound redundantly would be effective. For instance, it is possible that there was not a benefit of redundant designs for Experiment 1 and 2 because the dimensions of sounds used were perceived separately. If integral dimensions of sound are used redundantly in an auditory display, it may be that the same redundancy gains seen with speeded card sorting tasks (Garner, 1974) would generalize to a task such as the one presented here.

One must be careful when generalizing these results beyond the stimuli tested. The specific values of location, skew, and spread tested no doubt affected the relative difficulty of these dimensions. Moreover, further research is needed to see if these results hold on tasks with dependent variables other than the ability to match the auditory and visual displays. It is also an open question as to whether the finding that the temporal dimension is better than pitch would hold in other contexts. It is possible that extracting information from the temporal presentation is more attention demanding than from the pitch dimension. If this is the case, then the use of the temporal dimension in dual task situations where visual and auditory information must be interpreted simultaneously might result in poorer performance than would the pitch dimension. The Experiment 3 was designed to investigate these open issues.

Experiment 3

In the previous experiments, the task was for the participants to match an auditory box plot to a visual box plot. This was chosen because it is a straightforward task and easy to learn. However, some participants may have focused on pattern matching rather than understanding the data distribution represented by the box plot. Also, while a good starting point for investigating the sonification of box plots, the task in Experiments 1 and 2 is not one with much applicability to the real world. To address both of these issues, the task Experiment 3 was designed to be one step closer to a real task where individuals are having to monitor data and identify whether or not the data presented meet certain criteria. For example, someone might have the task of monitoring the pressure of the oil being pumped from an oil well. If the pressure is too high for too long, or unevenly distributed, it could result in damage to the oil platform. If an auditory display were used in this situation, the attendant could monitor the average pressure and how evenly the pressure values are distributed for an interval of time by listening to an auditory box plot representing the pressure measurements for that time interval. The attendant could identify if the means were too high and whether the pressure values were evenly distributed. The task used in Experiment 3 is similar in that it required participants to listened to auditory box plots and identify whether or not the box plot was out of range. If it was out of range, the participants indicated which parameter was "off target" (i.e., it did not meet the target specifications for central tendency or skew). To do this task successfully required participants to attend to the statistical parameters of the data.

An intriguing finding in the first experiment was the absence of a benefit in participants' performance when a redundant mapping was used. As mentioned previously, this pattern was consistent with the results of studies using a simple task in which participants map sounds to absolute numeric values (Sándor & Lane, 2003). Thus, there has been no benefit in performance with auditory designs that use two dimensions to redundantly represent data. To examine whether this generalizes to other situations, Experiment 3 was designed to replicate and extend the previous findings with a different task and with different dimensions of sound.

It is well known that some pairs of dimensions show facilitation with sorting tasks if they are presented redundantly whereas others do not. There is good evidence that redundancy gains are found for dimensions that are perceived integrally rather than as separate dimensions (Garner, 1974). Experiment 3 investigated whether these findings generalize to a monitoring task. Specifically, the study investigated whether the effect of redundant mapping result is different depending on whether the dimensions are integral or separable.

Previous research has found that frequency (or pitch) and loudness are integral (Grau & Kemler-Nelson, 1988; Neuhoff, McBeath, & Wanzie, 1999) whereas pitch and temporal are separable (Palmer & Krumhansl, 1987), although the findings for the latter are not as consistent as those for the former. Experiment 3 used these dimension pairs (pitch and loudness, pitch and temporal) in the design of the auditory box plots. Specifically, all of the participants had a condition in which redundant mappings are used for the sonified box plots. However, for one of the groups the redundant condition used the sound dimensions of pitch and loudness and for the other group pitch and time were the sound dimensions for the redundant condition. For each group, the dimensions were used both individually and redundantly to map the auditory box plots. For example, the participants in the pitch-loudness group performed the monitoring task in three different sound design conditions: pitch, loudness, and redundant (pitch and loudness).

One of the more valuable potential applications of data sonification is an "eyes busy" situation. These are situations requiring divided attention between a visual task and an auditory task simultaneously. Dual task performance generally places large demands on central capacity and results in decrements in task performance (Kahneman, 1973; Navon & Gopher, 1979; Norman & Bobrow, 1975). The investigation of these decrements is important to inform any application of sonification used concurrently with an eyes-busy task. There are many theories on how humans allocate resources in a divided attention task (e.g. rapid switching, performance resource functions, etc.), and many of the questions regarding these theories remained unresolved. While the research presented here is interested in examining the effects of divided attention on performance, it is not designed to specifically address the theories of divided attention. Instead, one purpose of the Experiment 3 was to investigate whether the finding that participants are more accurate with temporal than with pitch mapping would generalize to situations in which there is little available processing capacity. As mentioned previously, it may be that the temporal dimension is more attention demanding than the pitch dimension and would therefore not be as effective in a dual-task situation. To investigate this, the participants in Experiment 3 had a concurrent visual task for 50% of the trials. The visual task in the Experiment 3 consisted of a target detection task in which the participant saw a series of visual images and had to indicate, as quickly as possible, when a visual target was been presented. A visual task with a reaction time requirement was chosen because it is thought to be attention demanding. This type of task is also, in an abstract way, similar to a real monitoring task where an individual is attending to information on visual display while simultaneously monitoring information from an auditory source.

An interesting finding from Experiment 1 was the disassociation between preference and performance for the auditory designs. Most participants preferred the redundant design to a single mapping even though their performance was higher on a single mapping design (using pitch). However, in Experiment 2, participants seemed to have no preference for the different mappings and they thought the timing and redundant conditions were the least difficult. Given that these two studies do not show a clear pattern and previous research has not been conclusive regarding any dissociation between preference and performance on auditory tasks, Experiment 3 continued to address this issue. Specifically, participants were asked to rate the sounds and the tasks after they had completed the experiment. These results were examined to see if participants had different preference ratings for redundant designs using integral versus separable dimensions.

Method

Participants. All participants were undergraduate students between the ages of 18 and 24 from Rice University who received course credit and were randomly assigned to one of the two experimental groups – Integral or Separable. Initially, there were 66 participants, 32 in the Integral group and 34 in the Separable. However, the loudness mapping was originally done incorrectly and thus the data from the 32 participants in the Integral condition were not used. The mappings for the loudness and redundant conditions in the Integral were corrected and the experiment was run again for the Integral group. After all of the data were collected, there were 9 participants whose data was eliminated, 5 because of problems with the computer (1 in the separable condition and 4 in the integral condition) and 2 who reported hearing loss after they had started the experiment. This resulted in a total of 66 participants, 33 in the Integral group and 33 in the Separable group. There were 25 males and 41 females overall.

Procedure. Participants completed a 30-minute training session that included examples of the auditory box plots they heard during the experiment in addition to 30 trials of the auditory task. After the training, they did the experiment and then completed a survey asking them to rate the sound conditions and the auditory monitoring task. In addition to the subjective ratings, the survey asked for demographic data from the participants. The experiment lasted approximately one and a half hours and consisted of

300 trials. The trials were completed in blocks of 25 with each block having one sound condition presented. The participants did four sets of three blocks and the order that the sound conditions were presented was counter-balanced. Thus the participants would have 25 trials with one type of sound, 25 trials with another type of sound and then 25 trials with the third type of sound. The sequence would then repeat three more times.

Equipment. Macintosh computers with Internet access were used for the experiment. All of the computers were in the Sonification laboratory at Rice University. The training, experiment, and data collection were done using an interactive website with Java programming.

Design. The experiment was a factorial design with one between group variable (Integrality of dimension: 2 levels) and two within group variables (Auditory design – 3 levels; Concurrent task – 2 levels). Table 6 outlines the design of the study and shows that for each of the Integrality of dimension groups, the participants monitored box plots with three different auditory designs: two single dimension auditory designs and one redundant design. In addition to what is outlined in Table 6, for 50% of the trials, the participants were required to perform a visual monitoring task while they were conducting the auditory monitoring task. A visual target was presented in 20% of the visual monitoring trials. The participants were told that the visual task was the most important task and to focus their efforts on it.

Table 6
Outline of the between and within variables.

Integrality of dimension group (Between variable)	Auditory design (Within variable)		
Pitch/ Loudness	Pitch	Loudness	Redundant
Pitch/ Temporal	Pitch	Temporal	Redundant

Task. The auditory task in this experiment was a monitoring task where the participants listened to auditory box plots and identified when a box plot was "off target" for two different parameters (i.e., skew or location). The participants were instructed that the box plot could be off target for either skew or location but not both. After the listening to the auditory box plot, the participants indicated whether the box plot was on target, off target because of skew, or off target because of location. Participants were given feedback on their performance throughout the training and the experiment.

In addition to the monitoring task, for 50% of the trials, participants had a concurrent visual task to perform during the experiment. This task was a target detection task and the participants saw a series of images on the computer screen while they were listening to the box plots. When a target image appeared, the participants were instructed to respond as quickly as possible by using the mouse to click a button on the screen. Participants were given feedback on their performance. Specifically, if the participants indicated they saw a target image (by clicking on the button), the monitoring session would stop and they would be given feedback at that time. If they did not click on the button (indicating that they had not seen a target image), participants would be given feedback after the auditory monitoring session ended.

As mentioned previously, targets appeared on 20% of the trials and only the trials without a target were used for data analysis. Participants also had trials where they only performed the monitoring task (i.e., making judgments about the auditory box plots) thus, during these trials no images were presented.

For each trial of the auditory monitoring task, the auditory box plot was presented through headphones and after it had finished playing, three buttons would appear on the screen. The participants would click on one of the buttons to categorize the auditory box plot as “In-Control,” “Out-of-Control Location,” or “Out-of-Control Skew” and they were given feedback on their selection. If the participants incorrectly categorized the auditory box plot on the first attempts, they were instructed to try again until they got it correct. The auditory box plots were each 8.2 seconds long and the next trial would begin after the participant clicked a “Next Trial” button. For the visual task, 10 images were presented while the auditory box plot was playing, thus the image changed every .82 seconds. The images were 215 pixels by 223 pixels and were presented on an eMac computer with a 17-inch flat screen monitor. The images consisted of black circles with a diameter of 180 pixels on a gray background with a portion of the circle gone. Target images were those with 75% or more of the circle in black. Figure 3a shows an example of a target image and Figure 3b shows an example of a non-target image. Data on participants’ accuracy were collected for both the monitoring task and the visual task.

Participants were informed that the visual task simulates a monitoring scenario and, as mentioned before, the information represented by the images is the most critical.

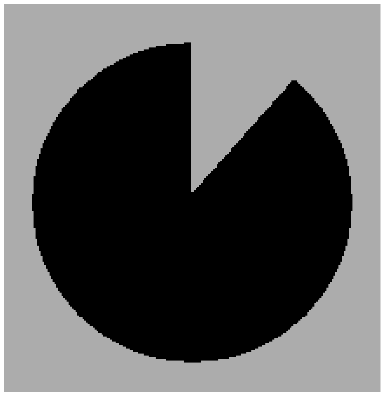

Figure 3a
Example of a target image

Figure 3b
Example of a non-target image (or a distracter)

Stimuli. There were 75 different box plots used in the study and the values of these box plots were based on scale from –16.0 to 16.0. The box plots were built by combing different levels of skewness and location (both in-control and out-of-control). Skewness was varied by manipulating the proportions between the two inner quartiles and the two outer quartiles and location was varied by manipulating the median. The "out-of-control location" box plots had a median ranging from –4.0 to –6.0 while the "in-control" box plots had a median ranging from 1.0 to –1.0.

One-third of the box plots were "in control" and were built by combining 5 levels of in-control skewness and in-control location. One-third of the box plots were out-of-control skew and were built using 5 levels of out-of-control skewness and in-control location. The last third of the box plots were out-of-control location and were built using 5 levels of out-of-control location and 5 levels of in-control skewness. The box plots were randomly selected for each trial from the set of 75 box plots.

For the separable dimensions, both auditory mapping methods were identical to the ones used in the experiments reported previously. For the pitch mapping, the values from the box plots were mapped to a note on the equal tempered scale in the range of 16 notes below and 16 notes above 440 Hz (middle C). For the temporal mapping, the distances between the values of the box plot were represented by the time between the onsets of the sounds and the pitch of all the sounds remained constant at 440 Hz. and the redundant condition was a combination of the pitch and temporal mapping.

For the integral dimensions, the pitch condition was identical to the one in the separable condition. The loudness condition used a range of 40 to 80 decibels and the values of the box plots were mapped to a decibel level in this range. For the redundant condition, the pitch and the loudness mappings were used and both were adjusted so that the perceived pitch and loudness of each of the redundant sounds would be the same as the perceived pitch or loudness in either of the single dimension mappings. The algorithms used to make these adjustments were obtained from the International Organization for Standardization ISO 226 "Acoustics - Normal equal-loudness level contours"(Nielsen & Brinkmann, 2003). The decibel and Hz levels used in Experiment 3 were within the ranges used by Grau & Nelson in their work that documents the integrality of pitch and loudness in a speed sorting task (Grau & Kemler-Nelson, 1988). Thus these ranges were assumed to be integral for the sake of stimuli here. All of the box plots had a duration of 8.2 seconds.

Measures. Performance on the visual task included the participants' accurately identifying the target image. For the auditory task, Hit and False Alarm rates were used to calculate d' for the "out-of-control skew" (Skew) and the "out-of-control location" (Location) box plots. These values were calculated because they allowed investigation of the participants' ability to discriminate the targets from the distracters and whether this differed by the independent variables - integrality of dimensions (integrality), auditory design (design), and task environment (task). For all of the results presented, the hit and false alarm rates will be presented in addition to the d'. These measures were investigated in summary to look at overall performance as well as across the four blocks to investigate the changes in performance over time.

Results

Visual Task. As can be seen in Table 7, performance on the visual task was very high and did not differ by integrality (integral vs. separable) or auditory design (pitch, temporal or loudness, or redundant).

Table 7
Proportion of correct responses for the visual task as a function of Integrality of Dimension and Auditory Design.

	Integral	Separable
Pitch	0.968	0.957
Loudness/Temporal	0.956	0.958
Redundant	0.958	0.965

Auditory task. Before any analysis of the auditory task was conducted, the following trials were excluded: any trials with a visual target presented (~10% of the trials), trials where the participant incorrectly responded that a target was present (228 out of 22,716), and trials where there was no record of any response from the participant (~.002% of the trials).

As mentioned previously, each participant had four blocks of the three types of auditory design. Overall, participants' performance increased over the four blocks (Table 8) and this increase did not differ by integrality, design, or task (see Appendix A – D for the specific hit and false alarm rates by integrality, design, and task). The first block was considered training and all subsequent analysis was conducted on the observations from block 2 – 4.

Table 8
Hit and False Alarm Rates as a function of Integrality and Block.

	Block 1	Block 2	Block 3	Block 4
Hits				
Integral	0.42	0.51	0.56	0.54
Separable	0.39	0.40	0.40	0.41
False Alarms				
Integral	0.50	0.47	0.57	0.58
Separable	0.49	0.61	0.52	0.74

As mentioned previously, the hit and false alarm rates were used to calculate the d' scores for each subject. For both the skewed and location box plots, there were several participants whose d' had to be estimated because it could not be calculated. For example, there were participants who had no false alarms and perfect hit rates for some conditions. A d' for these participants would be infinitely large and thus had to be approximated for the purposes of data analysis. Table 9 shows the types of situations where d' could not be calculated and the d' assigned to those participants for those conditions. There were 29 participants with a false alarm rate of zero, 21 with participants with a hit rate of 1 and 5 participants with a hit rate of 0.

Table 9
Criterion for assigning d' for situations when it could not be calculated.

Hit Rate	False Alarm Rate	Assigned d'
0.00	--	0.0
> .05	0	3.5
1.00	> 0	3.5
1.00	0	4.0

Table 10 and 11 gives the hit, false alarm, and d' rates for the Skewed and Location box plots, respectively. To investigate any benefit of a redundant design, for both of the integrality groups, the mean d' values of the better of the two single dimension were compared to the mean d' values for the redundant designs. As seen in Tables 10 and 11, this was pitch for the integral conditions and temporal for the separable conditions.

Table 10
Hit Rate, False Alarm Rate, and d' for Skew box plots.

Integral	Pitch	Loudness	Redundant
Hit Rate			
Single	0.63	0.54	0.69
Dual	0.60	0.42	0.60
False Alarm Rate			
Single	0.17	0.17	0.18
Dual	0.20	0.16	0.19
d'			
Single	1.59	1.29	1.89
Dual	1.38	1.17	1.71
Separable	**Pitch**	**Temporal**	**Redundant**
Hit Rate			
Single	0.57	0.75	0.75
Dual	0.53	0.70	0.77
False Alarm Rate			
Single	0.16	0.18	0.16
Dual	0.18	0.16	0.20
d'			
Single	1.45	1.92	1.95
Dual	1.17	2.10	1.99

Table 11
Hit Rate, False Alarm Rate, and d' for Location box plots.

Integral	Pitch	Loudness	Redundant
Hit Rate			
Single	0.41	0.51	0.51
Dual	0.41	0.42	0.47
False Alarm Rate			
Single	0.17	0.19	0.17
Dual	0.18	0.21	0.17
d'			
Single	0.85	0.97	1.14
Dual	0.88	0.74	1.17
Separable	**Pitch**	**Temporal**	**Redundant**
Hit Rate			
Single	0.46	0.57	0.61
Dual	0.35	0.53	0.53
False Alarm Rate			
Single	0.17	0.12	0.14
Dual	0.21	0.14	0.15
d'			
Single	0.98	1.58	1.70
Dual	0.70	1.44	1.32

An integrality (2: integral, separable), design (2:best, redundant), task (2: single task, dual task), and distribution type (2: Skew, Location) mixed design (integrality – between subject; design, task, and distribution type – within subject) ANOVA was conducted to investigate the effects of the independent variables on the participants' ability to identify the Skewed and Location distributions. Figure 4 shows the distributions of d' values for the integral and separable conditions. As seen in Figure 4a, for the integral conditions, the d' values are higher for the redundant design that the better design. However, in the separable conditions (Figure 4b) there is practically no difference between the better and the redundant designs. This design by integrality interaction was significant, $F(1,64) = 5.13$, $p = 0.027$, and did not differ as a function of distribution type or task, $F(1,64) = 0.09$, $p = 0.77$ and $F(1, 64) = 0.53$, $p = 0.470$ respectively. While the performance for the dual task conditions was lower overall, these differences are not significant nor were any of the interactions with stimuli, design, or integrality.

Although Block 1 was considered practice, it seemed worthwhile to investigate the effects of the independent variables on performance for this block. As seen in Table 12, participants had lower d' scores for the dual task than the single task, $F(1, 64) = 5.61$, $p = 0.021$. This difference was a substantial for the Location distributions and a smaller for the Skew distributions. This interaction was significant, $F(1, 64) = 14.35$, $p<0.001$. Neither the effect of integrality, $F(1, 64) = 0.25$, $p = 0.616$, or any other interaction approached significance.

Table 12
Mean d' for the first Block as a function of Task, Integrality, and Distribution type.

	Single	Dual	Difference
Location			
Integral	0.80	0.54	0.27
Separable	0.93	0.75	0.18
Skew			
Integral	1.17	1.02	0.15
Separable	1.18	1.08	0.09

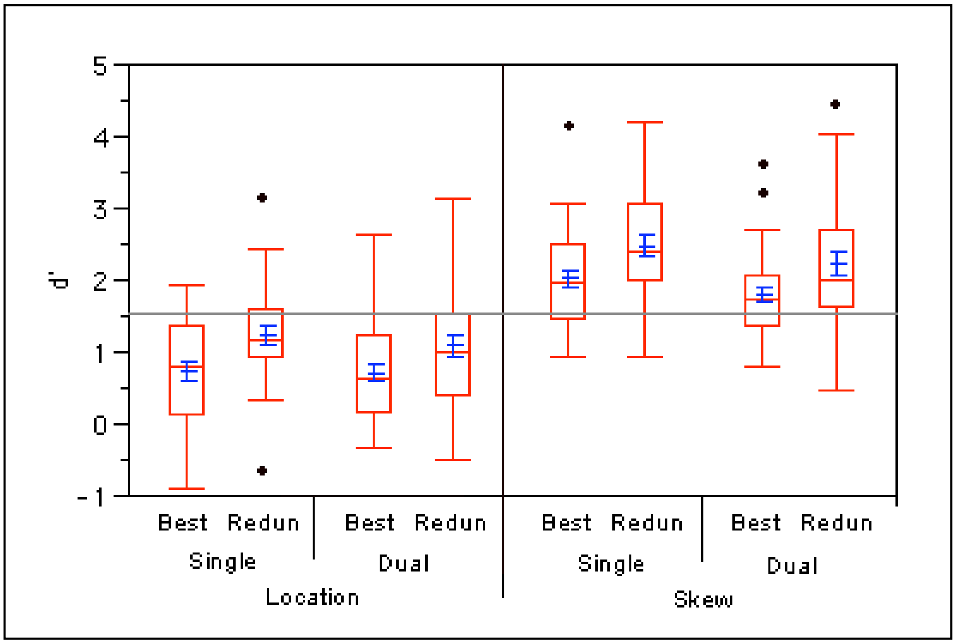

Figure 4a Box plots for the Integral conditions.

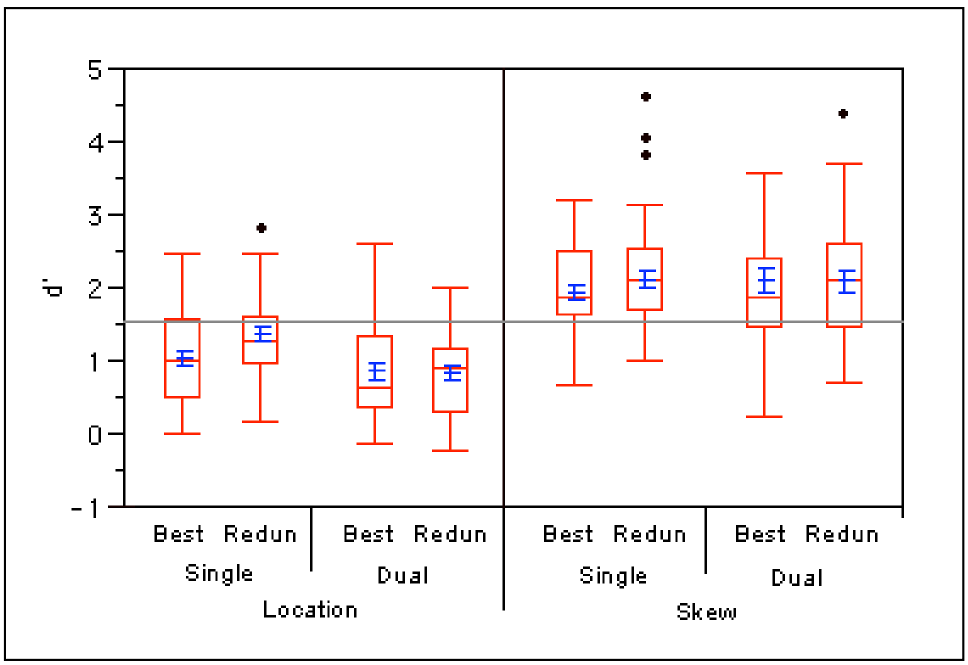

Figure 4b Box plots for the Separable conditions.

Figure 4 Box plots for the Integral (3a) and Separable (3b) conditions as a function of Distribution type, Task and Design. In addition to the median, each box plot shows the mean ± 1 standard error. The horizontal lines traversing the graphs indicate the grand means.

After the experiment was over, participants completed a survey asking them to rate the different sounds. They were asked to rate the difficulty, unpleasantness and annoyance of each of the sounds. A five point Likert scale was used for their responses and it ranged from 1 to 5 (e.g. for unpleasant; 1 - pleasant and 5 - unpleasant). Figures 5 and 6 show the mean responses for the integral and separable conditions. For the integral conditions, participants rated the redundant condition as less difficult, $F(2, 31) = 3.62$, $p = 0.038$, however they did not rate this condition as less annoying, $F(2,31) = 1.50$, $p = 0.238$ or less unpleasant, $F(2.31) = 0.26$, $p = 0.77$ respectively. For the separable conditions, the participants rated the temporal design as less difficult $F(2, 31) = 8.65$, $p = 0.001$, less unpleasant, $F(2, 31) = 6.45$, $p = 0.005$ and less annoying, $F(2, 31) = 7.03$, $p = 0.003$.

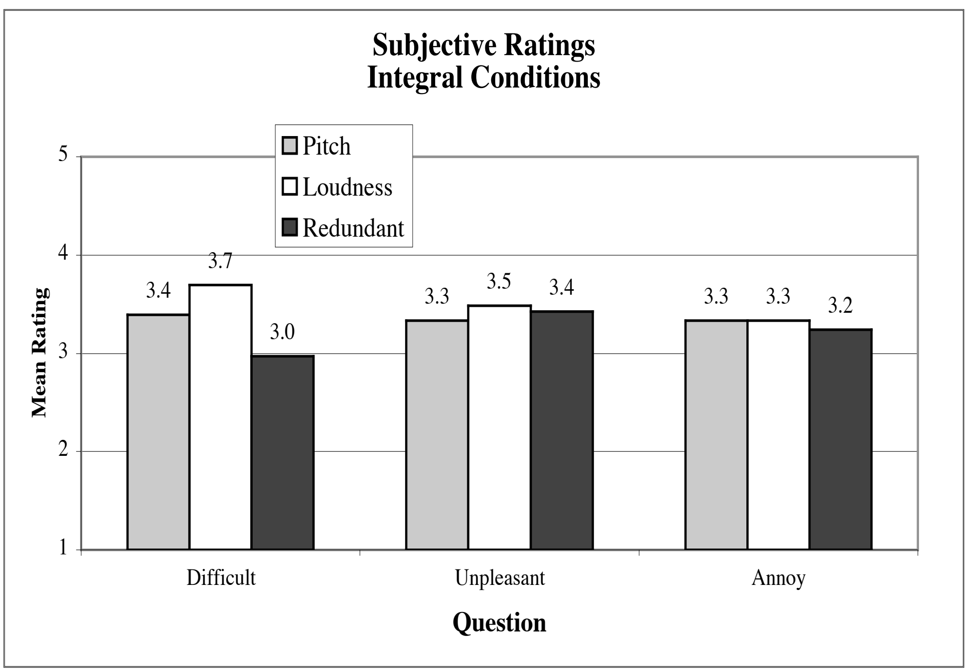

Figure 5 Mean ratings for the Integral conditions on questions asking participants' subjective impressions of the sounds

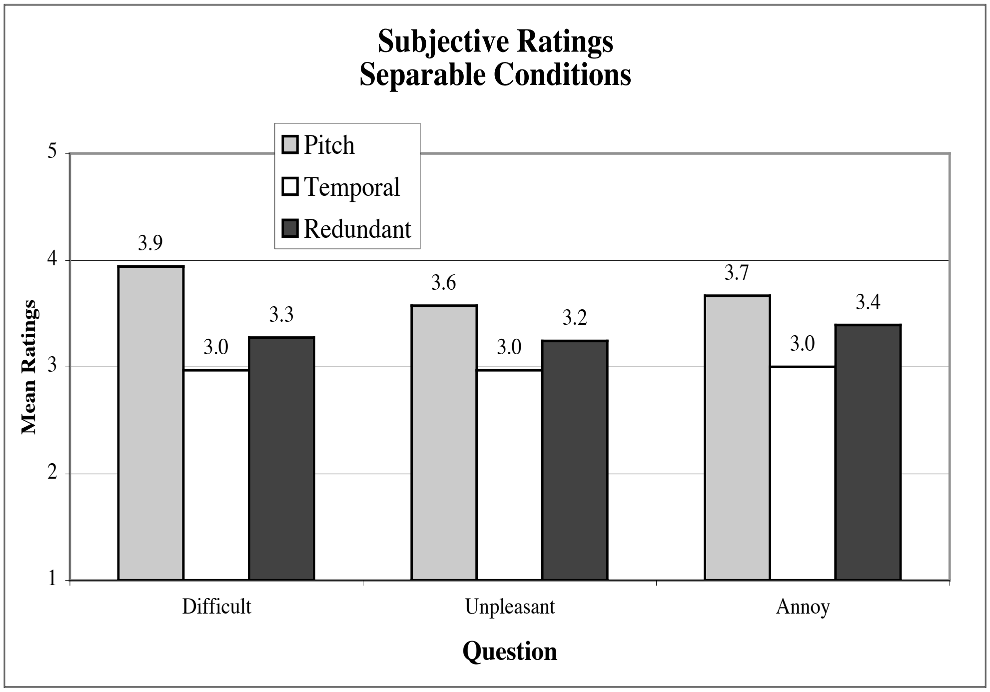

Figure 6 Mean ratings for the Separable conditions on questions asking participants' subjective impressions of the sounds

On the demographic survey, three of the participants indicated that they had perfect pitch. Often, people with perfect pitch perceive and interpret sounds differently than other people, so the data from these participants was analyzed separately. However, there were no systematic differences in these participants' performances.

Summary

One of the questions Experiment 3 was designed to address was whether the previous findings of no benefit for using dimensions of sound redundantly would generalize to auditory designs using integral dimensions of sound. For this task and these stimuli, the redundant design using integral dimensions of sound resulted in better performance over the better of the two single dimension mappings whereas for the separable dimensions, the redundant design did not benefit performance. The approximate difference in d’ (collapsing across task) between the best of the single dimensions and the redundant design for the Integral group was 0.25 for the Location distributions and 0.23 for the Skewed distributions. For the Separable group, these differences were not statistically significant and were only 0.01 for the Location distributions and 0.06 for the Skew distributions. These results suggest that for a task such as this, auditory displays using integral dimensions redundantly can improve the user’s performance.

The preference ratings for the different auditory displays were mixed. For the Integral group, participants perceived the redundant display as less difficult but they did not prefer it to either of the other two single dimensions. For the Separable group, where there was essentially no difference in their performance between the temporal and redundant design, participants preferred the temporal condition and perceived it as being less difficult.

The task in this experiment was designed to very roughly approximate a real-life monitoring situation where people would be simultaneously monitoring two sources of information. While performance was degraded for the dual task environment for the first block of trials, this effect did not continue after the first block. Participants may have learned to interleave the two tasks into one after the first block. This suggests that the use of auditory displays of data in an “eyes busy” environment is an appropriate application of the sonification of statistical data.

General Discussion

The redundancy gains found in the third experiment are consistent with the body of research on redundancy gains using a card-sorting task. However, these findings are inconsistent with those obtained by Sándor (2004) who used a much simpler task. Her research found no redundancy gain for integral dimensions in a task in which sounds were mapped to numeric values. The integral dimensions used in that research were duration and pitch as well as duration and rate of change in pitch. Both of these sets of dimensions were found to be integral (using a card sorting task), but there was no benefit of the use of these dimensions redundantly for her mapping task. It is an open question whether the differences in the findings are due to task differences or dimension differences and further research to address this is planned. It is also not certain that the redundancy effects occurred in the present experiment because the dimensions were integral. It could be simply that there is a redundancy effect for some pairs of dimensions and not others and it just so happened the two dimensions used here showed a redundancy gain.

Similar to the task in Experiments 1 and 2 there was a lot of variability in participants' performance with d's varying from –1.47 to 4.00. Participants generally found the task difficult. However, their performance improved after training and asymptoted quickly. Although Experiment 3 was not designed to compare participants' ability to identify skew or central tendency, it does seem that the skew was easier for the participants to identify. A possible explanation for this is that the number of values needed to identify a skewed box plots is fewer than the number needed to identify a location or control box plot. For the box plots presented in Experiment 3, all of the skewed distributions had three values close together (minimum, 25^{th} percentile and the median); thus, if three values were heard that were very similar, that display could easily be identified as skewed. For the "In-Control' and "Out-of-Control Location" stimuli, the participants had to listen to all five sounds in the auditory display before determining anything about the distribution represented by those sounds. It is also interesting to note that for Experiments 1 and 2, performance was better for the location box plots than the skewed box plot. One explanation for this is that the nature of the tasks are different enough that performance on the matching task did not generalize to the monitoring task.

One finding with implications for the value of the use of sonification in divided attention tasks was participants' ability (after 75 trials of practice) to perform equally well in the auditory monitoring task for both single and dual task environments. While there are some who feel that dual-task deficits almost always disappear with practice (Spelke, Hirst, & Neisser, 1976), the fact that so little practice was needed for this task is noteworthy for those looking to incorporate auditory displays into complicated monitoring environments.

The present research presented is consistent with other research that has found that the temporal dimension of sound is superior for auditory displays. The temporal dimension has been represented by pauses between the sounds, as in this experiment, as well as the duration of the sounds themselves (Sándor, 2004; Sándor & Lane, 2003; Watkins et al., 1992). The present results suggest that the superiority of the temporal dimension of sound can generalize to a more demanding monitoring environment like the one in Experiment 3 which required the participants to divided their attention as well as extract statistical information from the auditory box plots.

For all three experiments, the relationship between participants' performance and preferences is mixed with participants generally preferring the displays with a single temporal mapping even though this did not always result in the best performance. The findings that there was no preference for the redundant condition using integral dimensions is particularly interesting because of the substantially higher performance with this display than the displays using either the pitch or the loudness dimensions.

All of the experiments presented here used a fixed stimulus design and the different type of box plots were specifically designed to be distinguishable from each other. As with all designs of this type, it is possible that the results obtained here are limited to the specific stimuli used, thus caution must be used when generalizing these results.

The study of sonification often uses sound experimental methods to test the effects of particular displays and uses of these displays. However, the designs and implementations of these displays are often not well (or at all) guided by psychological theory. Appropriately applied, the theories of perception and information processing could provide important information and allow for predictions regarding how people will perform when interpreting auditory displays. Although these experiments were not designed to test a theoretical position, this paper presents results that suggest that the body of work on the perceptual effects of integral dimensions of sound (specifically, redundancy gains with integral dimensions) may be a theoretical starting place for investigating which and how many dimensions of sound to use in sonification.

References

Cleveland, W. (1994). *The elements of graphing data*. Murray Hill, New Jersey: AT&T Bell Laboratories.

Cleveland, W., & McGill, R. (1986). An experiment in graphical perception. *International Journal of Man-machine Studies, 25*, 491-500.

Flowers, J. H., Buhman, D. C., & 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., & Hauer, T. A. (1992). The ear's versus the eye's potential to access characteristics of numeric data. Are we too visuocentric? *Behavior Research Methods, Instruments and Computers, 24*(2), 258-264.

Flowers, J. H., & Hauer, T. A. (1993). "Sound" alternatives to visual graphics for exploratory data analysis. *Behavior Research Methods, Instruments and Computers, 25*(2), 242-249.

Flowers, J. H., & Hauer, T. A. (1995). Musical versus visual graphs: Cross-modal equivalence in perception of time series data. *Human Factors, 37*(3), 553-569.

Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.

Grau, J. W., & Kemler-Nelson, D. G. (1988). The distinction between integral and separable dimensions. Evidence for the integrality of pitch and loudness. *Journal of Experimental Psychology: General, 117*(4), 347-370.

Kahneman, D. (1973). *Attention and effort*. New Jersey: Englewood Cliffs.

Kramer, G. (1994). Some organizing principles for representing data with sound. In G. Kramer (Ed.), *Auditory display: Sonification, audification and auditory interfaces*. *Proceedings of the first international conference on auditory displays 1992* (pp. 185-222). Reading, MA: Addison-Wesley.

Kramer, G. (1996). *Mapping a single data stream to multiple auditory variables: A subjective approach to creating a compelling design*. Paper presented at the International Conference on Auditory Displays, Palo Alto, California, USA.

Lorho, G., Marila, J., & Hiipakka, J. (2001). *Feasibility of multiple non-speech sound presentations using headphones*. Paper presented at the International Conference on Auditory Displays, Espoo, Finland.

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review, 86*(3), 214-255.

Neuhoff, J., McBeath, M. K., & Wanzie, W. C. (1999). Dynamic frequency change influences loudness perception. A central, analytic process. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 1050-1051-1059.

Nielsen, L., & Brinkmann, K. (2003). *Normal equal-loudness-level contours*.: International Organization for Standardization.

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7*, 44-64.

Palmer, C., & Krumhansl, C. L. (1987). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance, 13*(1), 116-126.

Petrie, H., & Morley, S. (1998). *The use of non-speech sounds in non-visual interfaces to the MS-windows GUI for blind computer users*. Paper presented at the International Conference on Auditory Displays, Glasgow, UK.

Sándor, A. (2004). *Perceptual interaction of duration with pitch and rate of change in pitch: Implications for sonification*. Unpublished Master's, Rice University, Houston, Texas.

Sándor, A., & Lane, D. M. (2003). *Sonification of absolute values with single and multiple dimensions*. Paper presented at the International Conference on Auditory Displays, Boston, MA.

Spelke, E., Hirst, W., & Neisser, U. (1976). Skills of divided attention. *Cognition, 4*, 215-230.

Walker, B. N., & Lane, D. M. (2001). *Psychophysical scaling of sonification mappings: A comparison of visually impaired and sighted listeners*. Paper presented at the International Conference on Auditory Display, Espoo, Finland.

Watkins, M. J., LeCompte, D. C., Elliott, M. N., & Fish, S. B. (1992). Short-term memory for the timing of auditory and visual signals. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*(5), 931-937.

Appendices

Appendix A

Skew Hit Rate over time as a function of Integrality, Design and Task.

		Block 1	Block 2	Block 3	Block 4
Integral					
Loudness	Single	0.42	0.51	0.56	0.54
	Dual	0.39	0.40	0.40	0.41
Pitch	Single	0.52	0.60	0.66	0.62
	Dual	0.57	0.68	0.57	0.54
Redundant	Single	0.65	0.64	0.64	0.67
	Dual	0.47	0.59	0.55	0.54
Separable					
Pitch	Single	0.50	0.47	0.57	0.58
	Dual	0.49	0.61	0.52	0.74
Timing	Single	0.58	0.70	0.79	0.78
	Dual	0.55	0.68	0.70	0.67
Redundant	Single	0.67	0.68	0.79	0.81
	Dual	0.56	0.73	0.74	0.89

Appendix B

Skew False Alarm Rate over time as a function of Integrality, Design and Task.

		Block 1	Block 2	Block 3	Block 4
Integral					
Loudness	Single	0.21	0.19	0.17	0.17
	Dual	0.18	0.16	0.18	0.15
Pitch	Single	0.16	0.19	0.15	0.17
	Dual	0.23	0.16	0.22	0.21
Redundant	Single	0.17	0.19	0.17	0.17
	Dual	0.18	0.22	0.19	0.15
Separable					
Pitch	Single	0.19	0.15	0.18	0.13
	Dual	0.22	0.18	0.18	0.17
Timing	Single	0.25	0.24	0.19	0.13
	Dual	0.20	0.19	0.15	0.16
Redundant	Single	0.19	0.17	0.18	0.15
	Dual	0.21	0.20	0.24	0.17

Appendix C

Location Hit Rates over time as a function of Integrality, Design and Task.

		Block 1	Block 2	Block 3	Block 4
Integral					
Loudness	Single	0.37	0.43	0.37	0.50
	Dual	0.34	0.33	0.54	0.47
Pitch	Single	0.42	0.44	0.54	0.48
	Dual	0.36	0.51	0.43	0.44
Redundant	Single	0.50	0.44	0.53	0.57
	Dual	0.29	0.43	0.52	0.50
Separable					
Pitch	Single	0.44	0.50	0.48	0.46
	Dual	0.37	0.33	0.32	0.44
Timing	Single	0.49	0.55	0.55	0.56
	Dual	0.49	0.50	0.67	0.57
Redundant	Single	0.53	0.59	0.58	0.64
	Dual	0.50	0.60	0.64	0.62

Appendix D

Location False Alarm Rates over time as a function of Integrality, Design and Task.

		Block 1	Block 2	Block 3	Block 4
Integral					
Loudness	Single	0.19	0.18	0.21	0.18
	Dual	0.19	0.16	0.19	0.24
Pitch	Single	0.16	0.19	0.19	0.14
	Dual	0.19	0.16	0.18	0.18
Redundant	Single	0.19	0.15	0.16	0.19
	Dual	0.23	0.18	0.16	0.20
Separable					
Pitch	Single	0.25	0.20	0.16	0.15
	Dual	0.23	0.19	0.22	0.18
Timing	Single	0.21	0.16	0.10	0.11
	Dual	0.16	0.16	0.17	0.10
Redundant	Single	0.15	0.13	0.14	0.14
	Dual	0.17	0.17	0.16	0.12