Ecological Psychoacoustics and Auditory Displays: Hearing, Grouping, and Meaning Making

Bruce N. Walker and Gregory Kramer

1. Introduction

Auditory display is the use of non-speech sound to communicate information to a listener. The field of auditory display encompasses everything from alarms to complex sonifications. While there are many differences in the design of fire alarms and auditory graphs, the entire auditory display endeavor can benefit from and contribute to the study of ecological psychoacoustics.

A body of relevant knowledge about auditory perception and cognition is important to the successful design of auditory displays. We basically need to know enough about how people hear and think. Auditory display designers must then apply that knowledge, which is largely an issue of disseminating the information and training the designers to use it. The key question for this volume is whether traditional psychoacoustics research has provided the knowledge necessary to build useful and usable auditory displays. The short answer is that this research has provided a good starting point. Nevertheless, while a great deal of useful knowledge is available for basic perception of simple sounds, there remain many unaddressed scientific and practical questions that relate to the often complex sounds used in real-world auditory displays, the effects of typical listening environments, and the influence of the knowledge, experience, and expectations of the listener. A second, and perhaps equally important, question is how and what the field of auditory display research can contribute to psychoacoustics. We would contend that there is much to be shared in this direction as well.

One approach to this topic is to focus on the task of interacting with an auditory display to extract the meaning the designer intended. In this task-oriented approach, we can consider three general types of subtasks, each of which depends on the psychoacoustics research community for design recommendations. First, there is the "simple" perception of the sounds in the listening environment. If you cannot hear it or discern how it is changing, you cannot extract meaning from it. Many of the other chapters in this volume address particular aspects of this topic (e.g., Schmuckler;

Schauch). Second, there is the subtask of parsing the auditory scene into sound sources, or streams (i.e., segregating and grouping, Bregman, 1990, and see Van Valkenburg & Kubovy, this volume), and distinct variables such as pitch, loudness, and a host of timbral variable such as brightness or woodiness. Finally, there is the subtask of associative and cognitive processing that result in deriving meaning from the sounds in line with what the sound designer intended. All of these stages are required for successful communication of information via an auditory display.

The knowledge about each stage that has been contributed by the traditional psychophysical research community has been of considerable benefit to auditory displays that rely more on the first stage (perception), and has been somewhat less useful for the displays that place greater demands on the second and third stages (grouping and meaning-making). Our contention is that the more recent ecological form of psychoacoustics has much to offer all types of auditory display designs, especially those that attempt to convey more information or are employed in acoustically complex environments. In fact, there is an increasingly direct and synergistic relationship between the ecological research community and the auditory display community: as more is known about complex sound perception, auditory stream analysis, and the cognitive processing of sound in real-world situations, better and more sophisticated displays can be created; and as those more advanced displays are used, much can be learned about the auditory perception and cognition that is taking place.

In this chapter we present a brief history of auditory displays and establish a basic vocabulary for discussing the field. We then outline the three subtasks of basic perception, discerning streams and variables, and meaning making. Next, we examine the contributions of traditional psychoacoustic research to auditory display and where this research has fallen short of the needs of the field. We then examine where recent research into higher-level perceptual phenomena fills some of the gaps left by traditional research. We close with a short discussion of these research opportunities and how they stand to influence future auditory display design, and conversely how new complex auditory displays may contribute to the emerging field of ecological psychoacoustics.

2. Brief History and Terminology

Auditory display is a broad term referring to the use of any type of sound to present information to a listener. This may include, but is certainly not limited to, warnings, alarms, status indicators, and data sonification. Non-speech sounds have been used for a very long time to convey information. We provide only a brief introduction here, sufficient as a foundation for our subsequent discussion. See Kramer (1999) for an excellent summary of the types of auditory displays, and their development. Also, the amount of information that is intended to be conveyed by an auditory display is important to the topic of ecological psychoacoustics research. There is a wide range in the level of complexity of the intended message[1], but it has tended to increase over the years.

The simplest (not to mention earliest, and still most common) auditory displays are basic *alerts and notifications* (Sanders & McCormick, 1993; Sorkin, 1987). The sounds indicate that something has, or is about to happen, or that the listener's attention is required in some task. One example is the long beep indicating the cooking time on a microwave oven has expired. There is generally little information as to the details of the event—the microwave beep does not indicate if the food is cooked or not. Another commonly heard alert is the telephone ring—the basic ring tone does not indicate who is calling, or why. *Cautions and warnings* are alert sounds that specifically indicate one of a limited class of adverse events. In this case, the sounds communicate more about the nature of the event, by virtue of the unique association. Through experience and learning the listener comes to equate a certain wailing sound with a fire alarm. Again, the information about the events that the alarm conveys is usually limited to binary threshold-crossings. That is, the fire alarm signals that there is or is not a fire, not where or how hot it is. As technology and interfaces have advanced, there has been an increased

---

[1] Note that the focus here is on *intentional* sounds, namely those that are put into a system by a designer to convey specific information. *Incidental* sounds, such as the clicking and whirring of a hard drive, do provide status and activity information to a user, even though these sounds were certainly not engineered into the system. It is interesting and instructive to note that in many cases designers are now intentionally mimicking incidental sounds, in recognition of their utility. One example is the artificially introduced clicks that one often hears when "pressing" the software buttons on a touch screen interface.

need to provide more details about the events a sound is announcing. There have been two basic approaches to addressing this challenge: auditory icons and earcons.

Auditory icons are the auditory equivalent of visual icons, which are graphical symbols used to represent objects or processes. In the era of personal computing, icons have come to depict "functions of computer systems, the system objects, various types of system status, and [they] represent commands, cursors, menu items, windows, screen selection buttons, utilities, processes, programs and the like" (Uzilevsky & Andreev, 1993, p. 115). The ability of icons to simplify information display is largely due to their capacity to present a lot of information in a concise and easily recognized format (Blattner, Sumikawa, & Greenberg, 1989; Gittins, 1986), and this may be due to the visual system's ability to process several dimensions (shape, color, etc.) in parallel. In addition, icons are more easily located and processed than words (Hemenway, 1982; Shneiderman, 1998), and can help transcend linguistic and cultural barriers (Kolers, 1969; Ossner, 1990; Uzilevsky & Andreev, 1993).

Auditory icons, then, are brief sounds that have been introduced into the computer interface to represent objects, functions, and actions to the user via the auditory modality. Gaver suggests that, “Objects in the computer world should be represented by the objects involved in sound-producing events; actions by the interactions that cause sound; and attributes of the system environment... by attributes of the sonic environment” (Gaver, 1989, pp. 75-76). As an example, an auditory icon for a printer might be a brief sound of a line printer or typewriter. The directness or auditory similarity between the icon and the actual object can vary considerably. However, any sound that is intended to evoke the sound of an object or action is still generally classified as an auditory icon.

Modern interfaces involve many objects and actions for which there is no natural or even iconic sound, thereby precluding auditory icons. *Earcons* are sounds in an interface that represent a full range of “messages and functions, as well as states and labels” (Blattner et al., 1989). Earcons often employ a simple, hierarchical language of sounds and are often musical in nature. The relationship between the earcon and the action is, at most, *metaphorical*. An example is a three-note pattern representing a file, in which a decrease in loudness and pitch represents "file deletion"—the diminishing loudness and pitch of the sound is a metaphor for the destruction of the file. Most earcons have purely *symbolic* mappings between the sounds and the information they represent. They are arbitrary, and basically must be learned by the user. An example would be a plain "beep" to represent "file deletion," with no acoustic properties associated, even metaphorically, with the represented action. The iconic – metaphorical – symbolic distinction is not categorical so much as it represents a continuum of representational types (see Kramer, 1994b, for more on this topic).

While auditory icons and earcons use sound to represent an event more or less metaphorically, a different, and much more direct approach to conveying information with sound is to simply listen to the event, or listen to data generated by the event. *Audification* is the direct translation of a data waveform into sound (Kramer, 1994a). This often requires that the data wave be frequency-shifted into the audible range for humans, or time-shifted (slowed down or sped up) to allow for appropriate inspection by the listener. A common example of audification is the playback of seismic data. Because of their complexity, visual seismograms are difficult to understand and categorize. By speeding up the playback of the recorded signal so it falls into the human audible range, listeners have been able to classify seismic records as either atomic bomb blasts or earthquakes with accuracies greater than 90% (Speeth, 1961). More recent work has applied audification to oil exploration and further earthquake sonifications (Dombois, 2001, 2002; Hayward, 1994). Scarf (1979) used audification to explore data from the Voyager-2 space probe crossing the rings of Saturn. Visual data contained far too much noise for clear information to be extracted. However, certain types of wave structures had audio signatures that could easily be detected aurally, despite the background noise. Audification results demonstrate that some data are better dealt with by modalities other than the traditional visual mode.

Many extant auditory displays, other than caution and warning tones, have been used as dynamic *process monitors*, capitalizing on “the listener's ability to detect small changes in auditory events or the user's need to have their eyes free for other tasks” (Kramer et al., 1999, p. 3). Auditory displays have been developed for monitoring models of a cola bottling factory (Gaver, Smith, & O'Shea, 1991) and a crystal factory (B. N. Walker & Kramer, 1996b), among others, and for monitoring multiple streams of patient data in an anesthesiologist's workstation (Fitch & Kramer, 1994).

The more recent, and in general most sophisticated, auditory displays utilize several of the approaches mentioned to this point in order to allow a listener to extract meaning from the sound. *Sonification* is the use of non-speech audio to convey information such as that used in the interpretation of scientific results. Specifically, data sonification is “the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation” (Kramer et al., 1999, p. 3). That is, scientific data, of any sort, is used to change the parameters of a synthesized tone, usually in a largely metaphorical or symbolic manner (e.g., changing pitch represents changing rainfall or temperature, Flowers & Grafel, 2002). A well-known early example is the Geiger counter—radiation density is represented by the rate of clicks.

It is often helpful to think of recent sonification efforts as the process of creating sophisticated “auditory graphs,” analogous to the visualizations produced by modern graphing applications. Perhaps the first formal analysis of how to represent data using sound was Sara Bly's (1982) doctoral thesis from the University of California, Davis. Bly looked at how best to display multivariate, logarithmic, and time-varying data, using only sound. Part of sonification’s growing appeal is that it can be used to display highly complex and multidimensional data. Participants in Fitch and Kramer’s (1994) auditory anesthesiologist’s workstation monitored eight simultaneously changing patient variables, and performed significantly better with the auditory version than the visual display. Kramer has also described the sonification of five-dimensional financial data (Kramer, 1993) and nine dimensional chaotic data (Kramer & Ellison, 1991).

Recent studies have pointed out that successful interpretation of the sonified data requires more than just the sonified data—context is also necessary. This includes the auditory equivalents of axes, tick marks, trend lines, labels, and so on (Smith & Walker, 2002; B. N. Walker, 2000, 2002). The implementation of the context elements basically involves all of the auditory display concepts discussed to this point (notifications, warnings, etc.), as well as some that are particular to the auditory data interpretation task.

Recent research has used sonification to detect tumors in a medical application (Martins & Rangayyan, 1997), make discoveries in physics (Pereverzev, Loshak, Backhaus, Davis, & Packard, 1997), and analyze structural data from concrete highway bridges (Valenzuela, Sansalone, Krumhansl, & Streett, 1997). Another important area of application is in virtual environments, where denoting objects, validating user interactions, providing environmental context, and enhancing veridicality are all aided by the careful use of sound. It is quickly becoming clear that we are only scratching the surface of possibilities for using sound to display and help interpret complex scientific data. The key to continued utility is sorting out the more complex task elements that set modern sonification-style auditory displays apart from the somewhat-more-understood cautions and warnings.

By its very nature, sonification is interdisciplinary, integrating concepts from human perception, acoustics, design, the arts, and engineering. Thus, development of effective auditory representations of data requires interdisciplinary collaborations using the combined knowledge and efforts of psychologists, computer scientists, engineers, physicists, composers, and musicians, along with the expertise of specialists in the application areas being addressed. The potential contribution of ecological psychoacoustics is large and important to the continued evolution of auditory displays.

3. What Designers Need to Know: Auditory Display Subtasks

Auditory display can be understood as a form of applied auditory perception. Designers need to take into account a lot of acoustic, environmental, and human factors when creating an auditory display. As discussed at the outset, it is helpful to consider the various subtasks involved in the use of auditory displays. This approach highlights the needs of the auditory display design community, and frames our discussion of the successes and limits of traditional psychoacoustics research.

a) Perception

For all auditory displays the task of understanding the message is dominated by the perception subtask. That is, the most important aspects in their design relate to whether the listener can hear the sound and particular changes of particular parameters. This is especially true for more iconic displays such as notifications and warnings. It is crucial that the designer be able to predict if, when, and how well their sounds will be heard. Perception depends on a complex and dynamic interplay between the sound attributes, the environment, and the listener. For a design to be successful, issues of absolute detection thresholds, discrimination sensitivities, and masking top the list of issues an auditory display designer needs to consider. In order to have the optimal utility for designers, these factors need to be investigated using auditory display-like sounds that are presented in environments that are acoustically and cognitively realistic.

b) Stream Analysis

There are two main situations where streaming issues are particularly relevant to auditory display. The first is when streaming is related to detecting the correct auditory signal. For example, many auditory displays will be heard in an environment with multiple competing sound sources. Following the task-relevant source can be crucial. Sometimes the environment will be predictably "sourcy", such as in cockpit where there are multiple concurrent speech and non-speech auditory streams. In other circumstances, the designer will not know the exact number of competing sources, and will have to design for the worst-case scenario. The realm of computer interface notifications is a current example of where the designer must ensure distinction in the sound, in order for the listener to determine where the sound originated (e.g., to know if it was the computer that was beeping, as opposed to the PDA that is also sitting on the desktop).

The second main type of stream-dependence in auditory displays is when a designer intentionally maps informational distinctions onto different streams. One example is the spatial segregation of different data streams. The monthly sales figures from a company's Los Angeles office could be presented in the front-left auditory space, while at the same time the sales figures from the New York office could be presented in the front-right auditory space. This spatial separation should allow the user to mentally separate the two data sets. Further, the LA:west:left versus NY:east:right mapping is spatially compatible (c.f. Proctor & Reeve, 1990; B. N. Walker & Ehrenstein, 2000), which should help the listener remember which stream is which. Within a stream (say, the LA data), the sales data could be represented by the pitch of a continuously varying tone, while a series of intermittent clicks marks the passage of the days of the month. This would effectively result in two "substreams", one for data and one for context.

Thus, in order to make perceptible displays in a sourcy environment, or make effective use of streams in a sonification, it is important to know what acoustic characteristics allow for successful inter- and intra-stream segregation as well as grouping. It is also important to know how well listeners will be able to listen analytically to one stream (or substream) at some times, and listen holistically (integrating across streams) at another time. Again, this information needs to be available for ecologically valid sounds, environments, and listeners.

c) Conception and Meaning-making

Once the auditory display is heard and streamed into sources, the listener needs to understand the meaning that the sounds convey. The simplest type of meaning making is identification of one of a set of sounds. The acoustic attributes of a sound that make it distinct and memorable need to be applied to the design of alerts and warnings, to make sure that there is no ambiguity as to what the sound signifies. For example, the International Space Station caution and warning tones need to communicate unambiguously whether there is a fire, depressurization, or toxic environment condition, as all require immediate but different actions. It is important to know just how the manipulation of acoustic attributes such as the frequency or repetition rate affects a listener's categorization of the sound. For example, will a high-pitched alarm be perceived as more indicative of a fire or a depressurization?

A refinement of this is the study of the attributes that can allow a sound to be distinct and at the same time still belong within a family of sounds. This is important for the creation of metaphorical languages of sounds, as in the case of earcons. The computer object-related sounds (e.g., files, directories) need to be similar as a group, but different from the action-related sounds (e.g., edit, delete). Of course, the semantic rules used to create the earcons play a large role in this meaning making. The extent to which listeners are able to learn the "language" is also highly relevant.

Another type of cognitive interpretation is the representation of a class of information (i.e., data dimension) by a varying sound parameter (i.e., display dimension). A sonification example would be representing changing temperature by varying a sound's frequency. As Walker has summarized (2000; 2002), it is important for a display designer to know: (1) the optimal mapping: what sound parameter is best suited to represent a given type of data; (2) the optimal polarity: whether an increase in the sound parameter should indicate an increase or decrease in the data dimension; and (3) the optimal scaling function: how much change in a sound will represent a given change in the corresponding data dimension. Again, this information needs to be available for the kinds of sounds, environments, and listeners that will interact with a given auditory display.

4. What has traditional psychoacoustics contributed to auditory display?

A rich history of research has provided valuable insight into the physiological, perceptual, and cognitive aspects of auditory perception for relatively simple auditory events, such as pure tones and noise bursts. Much of this work has contributed to a functional knowledge base of various auditory thresholds, psychophysical scales, and models of auditory perception. Following on from the previous section, this can be organized by the perceptual dimensions that are important in auditory display, starting with those that contribute most to our understanding of the perception of sounds, then moving through what we know about streaming, and on to the research that indicates how a listener in a given context will interpret an auditory display. Other chapters in this volume cover many of the following areas much more thoroughly. To avoid redundancy, where possible we will refer the reader to those chapters for the details, restricting our coverage to the relevance to auditory displays.

a) A Framework: Vocabulary, Methods, and Limits

We should start by making it clear that the field of psychoacoustics has provided a foundation for the methodical evaluation of auditory displays and thus a link to existing scientific literature. Auditory display is, in large measure, an “engineering art”. As a solutions-driven endeavor, it requires guidelines and dependable results, and a language with which to communicate. The vocabulary of limits, just noticeable differences (JNDs), octave bands, fundamentals and harmonics, and auditory streams are all foundations for auditory display designers, as are descriptions of the listening environment, such as characteristics of background noise. In addition, the ways of defining sounds, in terms of frequency, decibels, and so on has had a large benefit to auditory display. (The musical vocabulary has also contributed enormously to the field of Auditory Display, but that is a separate topic, somewhat out of the scope of the present discussion.)

b) Rigorous experimental methods

The field of psychoacoustics has made available to the auditory display community a whole set of experimental and analysis approaches. These range from magnitude estimation procedures (e.g., Stevens, 1975) to assess perception of sound dimensions, to multidimensional scaling (e.g., Schiffman, Reynolds, & Young, 1981) used to assess the interaction of sound dimensions. More recently, more rigorous assessment of performance and effectiveness of auditory displays has involved experimental methods and statistical analysis techniques that, while not necessarily developed by psychoacoustics researchers, are characteristic of the methodical and quantitative heritage of psychoacoustics. The magnitude estimation literature (e.g., Stevens, 1975) has provided functions linking actual physical characteristics of the sound (amplitude, frequency, spectrum) with perceptions of those characteristics (e.g., loudness, pitch, timbre). This allows auditory displays to employ equal steps in the perceptual dimension, as opposed to the physical dimension, which is crucial for effective auditory displays.

c) Loudness

The most common attribute of sound used to convey information to a listener is its presence or absence. Warnings, status messages, and alerts are usually silent until a message needs to be communicated. It is crucial to know that the listener can detect the sound when required. Thus, the psychoacoustic research on minimum thresholds is important (e.g., see Hartmann, 1997; Licklider, 1951). Beyond the absolute detection of sounds, psychoacoustics has determined a power function that describes the way listeners perceive changes in sound intensity. This can allow a designer to create linear steps in loudness (i.e., perceived intensity), which may be more suited to representing certain types of data. At the upper end of the loudness scale, psychoacoustics has established guidelines for maximum exposure, both for single-event and continuous sounds (OSHA, 1981; Salvendy, 1997). This information allows a designer to create an effective and non-damaging auditory display, be it an alert, a fire alarm, or a cockpit sonification. See Schlauch, this volume, for a more thorough treatment of loudness.

d) Masking

The absolute limits of detecting a sound in ideal conditions (typical thresholds approach 20 microPascals SPL) are less important to the auditory display designer than determining if a particular sound will be heard in a given situation. The concept of masking, and the experiments that have determined the intensity, frequency, and temporal characteristics of masking have been helpful (for an overview, see Moore, 1989). Now designers can predict in advance if the sounds they design will be heard. There are effective masking models that take into account the properties of human hearing and allow for accurate assessment of competing sounds in a display, or the interplay of the display and background noise sources (e.g., Zwicker & Scharf, 1965). Patterson's (1982) work to study the masking of cockpit warning sounds was an example of the interplay of psychoacoustics research and auditory display design. Those research findings, and the guidelines that resulted, are now part of the auditory display vernacular.

e) Pitch

Pitch is the most commonly used auditory display dimension. This is because it is easy to manipulate and, generally speaking, changes in pitch are easily perceived. The fact that pitch is less influenced by the listening environment than is loudness makes it a more robust display dimension. The limits of frequency perception, often quoted at 20-20,000 Hz for young healthy listeners, are well known among the auditory display field. Widespread study of these limits provides a general context for auditory displays, so the designer knows to avoid frequencies much below 80 Hz, or above 10,000 Hz. However, perhaps even more important and useful is the experimental finding that sensitivity to pitch change varies with frequency (Robinson & Dadson, 1956), with peak sensitivity in the 3000 Hz range. Auditory display designers use this to center their displays in the most effectively perceived frequency region, typically 200-5000Hz. This also happens to correspond to the central range of typical musical instruments, which facilitates the use of MIDI and other music-based auditory display tools. Like masking concepts, research in the perception of harmonics can be used by auditory display designers to know in advance what pitches will be perceived, even if there are multiple frequencies in the display. The missing fundamental (Licklider, 1956; Schouten, 1940) is an example where psychoacoustic research helps auditory display developers to predict the perceptual experience of listeners in the pitch domain. Another area where the study of pitch perception has influenced auditory display is in the perception of slightly different frequencies. The beats that are heard when two sounds are detuned slightly can be used as a diagnostic tool in an auditory display, indicating that two parts are not aligned, or a process parameter is just off target (Wenzel, 1994). See Schmuckler, this volume, for a more thorough discussion of pitch perception in general.

f) Interacting Dimensions

Some perceptual dimensions are separable, and some interact (e.g., Garner, 1974). The knowledge that pitch and loudness are interacting dimensions is important for auditory displays. The availability of equal-loudness contours (ISO, 1987), based on recent efforts to refine the original work of Fletcher and Munson (1933) and Robinson and Dadson (1956), means that auditory displays can account for the non-linearities in human auditory perception. This makes auditory displays that rely on detection of a sound (e.g., warnings) more effective, and those that rely on the level of pitch or loudness more accurate. For example, a designer can use the equal-loudness contours, now described by a family of functions (ISO, 1987), to account for the effects of frequency on loudness. That is, an auditory display that uses frequency to represent temperature can be corrected, so that all of the sounds have equal-sounding loudness, thereby isolating the information to the frequency dimension (for an example of this technique in use see B. N. Walker, 2000, 2002). See, also, Neuhoff's chapter on sound interaction, this volume, for more on this topic.

g) Tempo and Rhythm

Tempo has been used in auditory displays for some time, especially to indicate the rate of a process or the frequency of an event (e.g., Geiger counter). From the psychoacoustics literature two basic features of auditory perception have been discovered that indicate sound can be effective for representing data in a variety of settings. First, auditory perception is particularly sensitive to temporal characteristics, or changes in sounds over time. Human auditory perception is well designed to discriminate between periodic and aperiodic events and can detect small changes in the temporal frequency of continuous signals (Fletcher, 1940; Resnick & Feth, 1975). This points to a distinct advantage of auditory over visual displays. Fast-changing or transient data that might be blurred or completely missed by visual displays may be easily detectable in even a primitive but well-designed auditory display (Kramer, 1994b). Thus, sonification is likely useful for comprehending or monitoring complex temporal data, or data that is embedded in other, more static, signals. Jones, this volume, presents more on timing and temporal attention.

h) Timbre

Timbre is a catch-all term in both psychoacoustics and auditory display, often used to mean all those sound attributes that are not loudness, pitch or tempo. The ability to distinguish sounds of different timbres has been important in mapping data to sounds. This is largely because timbre plays a large role in stream grouping and segregation. This allows, for example, one data stream played with a flute-like sound to be distinguished from a second data stream played with a clarinet-like sound. On the other side of the same coin, research on timbre has shown how sounds in displays can be similar (Ballas, 1993; Bonebright, 2001), facilitating the creation of sound families and hierarchies.

Also, the psychoacoustics (and plain acoustical physics) involved in sorting out what makes a particular instrument or other acoustical event, such as impact or scraping, sound like it does has been important in synthesizing sounds. Auditory icons rely on the timbral (and other) attributes of interface sounds to be representative of actual event-related sounds.

It should not be forgotten that aesthetics and acceptability plays a major role in the success of an auditory display. Timbre and overall spectral attributes are crucial in getting a display to sound correct and sound good. Wakefield, this volume, expands on timbre.

i) Pain and fatigue

Generally, designers avoid painful displays (!). However, without the psychophysical testing literature on both frequency and loudness (and their interaction), the limits of auditory pain would not be readily available to the auditory display designer. We can now simply ensure that sound levels and frequencies fall well within appropriate guideline limits (OSHA, 1981) (Salvendy, 1997). Much as with pain, the limits of using a system can be predicted quite well based not only on psychoacoustic experimentation designed to test fatigue and attention, but also from extensive serendipitous knowledge gained by conducting all sorts of listening experiments. Psychoacoustic researchers have found that listeners simply do not provide the same level of attention after half an hour of testing. These factors can be used by auditory display designers to maximize the effectiveness and acceptability of their systems, by avoiding the thresholds of auditory pain and fatigue.

j) Spatial Location

Psychoacoustic research on spatial hearing has been widely employed by auditory display developers. The extensive research in recent years on first understanding spatial perception of sound, and then developing methods for creating artificially spatialized sound (Wenzel, 1994) has opened the door for more sophisticated and effective auditory displays. As discussed earlier, using the spatial location of a sound as a display dimension can provide information in itself (location of a target), and can improve the number of channels or sources of data that can be displayed, especially in multitalker radio environments (Brungart, Ericson, & Simpson, 2002). The location of the sound can be used to separate (or group) different streams. Location can also analogously represent a change in a data variable. Psychoacoustics research has determined the angular resolution of our auditory system with both real and virtual sounds (approximately 1-2 degrees of azimuth in the front, and 5-6 degrees to the side, McKinley & Ericson, 1992; Mills, 1958), which is crucial if we want to know how well our display users will be able to estimate the data being represented (see also Carlile, Leong, & Hyams, 1997; Middlebrooks & Green, 1991).

Another useful aspect of spatial location that has been highlighted by psychoacoustic research is the fact that, unlike vision, perception of sound does not require the listener to be oriented in a particular direction. Auditory displays can therefore be used in situations where the eyes are already busy with another task. These characteristics make sound highly suitable for monitoring and alarm applications, particularly when these alarms may arise from many possible locations, or when visual attention may be diverted from the alarm location.

k) Streaming

Much is known about what general sound attributes will encourage or discourage grouping (see Bregman, 1990, for a thorough treatment). Perceptual research about streaming is very useful in an auditory display, where it can be used to cause different data representations to hang together, or remain distinct. Several examples have already been presented here. To examine one of the examples in a bit more detail, consider the use of streaming in providing context in an auditory graph (e.g., Smith & Walker, 2002). A steadily beeping sound can represent the maximum value in a data set, thus providing context, much like a grid line drawn on a visual graph. Making sure that the tempo of the "max" sound was quicker than that of the sound representing the data, resulted in two separate but coexistent auditory streams. For the most part, it was simply the *concept* of streams that was most useful to auditory display. The idea that there were ways to have multiple "auditory elements" in a display opened the door to high-dimensionality in auditory displays (e.g., up to 10 dimensions in Kramer, 1993). This possibility for auditory displays will continue to be one of its best selling features. Unfortunately, until recently, psychoacoustics research has not been able to provide much more than sets of guidelines that recommend ways to encourage segregation or grouping; there were certainly no hard and fast rules. However, see Van Valkenburg and Kubovy, this volume, for the latest developments that shed new light on the auditory scene.

l) Other aspects of perception

Other findings in the area of auditory perception bear on the promise of sound as a medium for data display and have helped to illuminate the optimal means of mapping data to specific dimensions of sound. These aspects include some studies on parallel listening (ability to monitor and process multiple auditory data sets, Astheimer, 1993), rapid detection, especially in high-stress environments (Mowbry & Gebhard, 1961), affective responses such as ease of learning and high engagement qualities of different sounds, and auditory gestalt formation and the discerning of relationships or trends in data streams (Kramer, 1994a; Kramer et al., 1999). McAdams and Bigand (1993) also present a collection of papers discussing various aspects of auditory cognition. All of these findings address somewhat more complex listening phenomena, and build on previous psychoacoustic knowledge in a seeming evolution toward studies of more dynamic and higher-level perception. All of these results can be used in the creation of more effective auditory displays, especially the results that provide more insight into the cognitive and meaning-extraction process in auditory display use.

m) Training and practice effects

One more research focus germane to the present discussion is the role of learning in auditory display efficacy. There are applications for which training is necessary to provide highly efficient performance. Blind users of assistive technology typically attain high skill levels, but only after extended practice (Earl & Leventhal, 1999). The special abilities of skilled sonar operators is another example that shows how learning can significantly enhance the efficiency with which auditory patterns can be discerned (Howard & Ballas, 1982). Since learning results in basic perception and higher-level cognitive processes becoming streamlined, a great amount of the cognitive psychology literature can also be useful for auditory display designers.

5. What have been the limitations of traditional psychoacoustics (in terms of auditory displays)?

a) Sound attributes

The most common limitation in the application of traditional psychoacoustics to auditory display research is the issue of *external validity,* that is, whether the experimental findings apply directly to real-world situations. First of all, the nature of the sounds in psychoacoustics research is typically unlike those in real applications. The test sounds used in traditional research are usually simpler (often sine waves), cleaner, clearer, and presented at a higher signal-to-noise ratio than is typically available for auditory displays. When sound parameters are changed in the course of an experiment, it is generally only one, or sometimes two parameters that are changed, while the rest are held constant. Further, within a trial or listening episode the sounds are usually static, unlike many auditory display sounds that vary continuously and often rapidly, in terms of pitch, tempo, timbre, or volume as data values change. That is, in a psychoacoustic experiment, pitch might be different from stimulus to stimulus, or trial to trial, but in an auditory display the sounds are often changing continuously. Walker and Ehrenstein (2000) showed that these *changes* in pitch, and not only the pitch value itself, can affect perception and performance with auditory displays.

In addition to the dynamics of sounds, the spectral characteristics (timbre) of laboratory sounds are often unlike those in real applications. Auditory display sounds that need to be realistic, or even just iconically related to real sounds, will have rich, complex spectra. The detection, perception, and comprehension of these sounds may very well be different from that of the simple tones typical of psychoacoustics experiments. This may be especially relevant in situations where the complex sounds coincide with other complex sounds, where issues of masking, beats, hidden frequencies, and so on may play out differently. It should be noted that it is not certain that major differences would be apparent—it is just that the results are not generally available, so the auditory display designer at present has no choice but to extrapolate from what is known about the perception and interpretation of simpler and more static sounds.

b) Listening environment

Typical psychoacoustics research is regularly conducted in some form of “sound attenuated” listening environment, often involving an isolated booth or even an anechoic chamber. This leads to a high signal-to-noise ratio, and removes virtually all competing sounds (not to mention most or all echoes). However, this is not at all typical of real auditory display listening situations, especially in military or industrial applications such as in-vehicle communication, assembly line warnings, and so on. Even a regular office can have 40-60 dB of background noise from sources that include computers and equipment, other workers, telephones ringing, and so on. This limits the direct applicability of psychoacoustic findings about perception of sounds. The two main problems with this are (1) the masking issues that arise in the real listening environment, and (2) issues related to attention and meaning-making. Any competing sounds that are present in a psychoacoustics listening study are usually introduced by the experimenter, and are often very unlike real competing sounds. Noise of various specific types (e.g., white, pink, brown) is generally used to simulate external noises. In some cases, real-world distracters (like airplane engine roar) can be successfully simulated by filtered artificial noise, however most competing sounds are not so predictable.

The other issue here is how competing sounds affect comprehension. It is well known that divided attention can seriously affect perception (e.g., Egan, Carterette, & Thwing, 1954; Treisman, 1964, and see Jones, this volume), but an additional and more cognitive problem arises when there are different and incompatible sonifications competing. The designer of an auditory graphing application for sales data might represent increasing dollar values with increasing pitch. However, another auditory display application in the operating system (or elsewhere in the environment) might also use increasing pitch to represent the progress in a large file transfer. The designers likely never anticipated that their systems would compete with another system for the attentional resources of the listener (not to mention the sound output capabilities of the computer's sound card). Thus, the environment can influence both simple perception and more cognitive elements of the task, so researchers need to investigate this in the context of auditory displays.

c) How parameters interact within a sound

Pollack and Ficks (1954) found that displays using multiple parameters of sound generally outperformed unidimensional displays measured elsewhere. Their research also indicated that the subdivision of display dimensions into finer levels does not improve information transmission as much as increasing the number of display dimensions. Kramer (1996) has suggested multiple mappings of data, whereby one series of data is used to change two or more parameters of the sound. For example, temperature might drive pitch, tempo, and brightness. How the resulting auditory display is perceived, and the meaning that would be communicated, remains to be tested rigorously. We should point out the work of Hellier, Edworthy, and their colleagues, where the perceived urgency of an alarm was varied by adjusting the sound attributes (e.g., Edworthy, Loxley, & Dennis, 1991; see also Guillaume, Drake, Rivenez, Pellieux, & Chastres, 2002; Hellier, Edworthy, & Dennis, 1993). Adjusting more than one sound attribute (e.g., frequency and amplitude envelope) had a greater effect than achieved by adjusting just one attribute. All this leads to clear (and still unanswered) questions about which sound dimensions to involve in a display, and how to use them alone or in combination to evoke specific perceptual or conceptual responses.

d) Limitations in tasks and cognitive demands

The tasks participants are asked to do in traditional psychoacoustical research are usually highly contrived and possess limited external validity. There are often very few cognitive requirements in the task. For example, detecting a signal sound amidst a background masker is instructive for listening researchers, but is of narrow utility to auditory display research. A real-world detection task would likely involve several other cognitive elements, often including interpreting the meaning in the sounds and then deciding on an appropriate course of action as a result. Cognitive load and distraction are typically achieved through added sounds or secondary tasks like remembering a number or counting backward, which have little to do with the complex perception and cognition tasks involved in auditory display perception. Simple alerts and warnings might be less affected by task demand increases, but, as pointed out by Smith and Walker (2002, p. 365), even the "basic" task of determining the price of a stock at a given time of day can involve the following perceptual and cognitive subtasks: "When given the opening [stock] price and asked to report the price at a given time (noon, for example), the subject must listen to the entire graph, recall the pitch he or she perceived at approximately half the duration (noon time), compare it to the pitch perceived at the very onset of the graph (the opening of trading), estimate the change in price represented by the difference between the noon-time pitch relative to the opening pitch, and add or subtract that change in price to the opening price of the stock." The cognitive demands of traditional psychoacoustics experiments do not even approach this level of complexity.

e) Metaphors and conceptual mappings

The relationships between underlying data and their acoustical representations need not be arbitrary. At the metaphorical and symbolic end of the auditory display spectrum, data are represented by sounds that often have little or no apparent relation to the sound. For example, pitch may be used to represent the volume of product sales; this is an arbitrary and purely analogic mapping. It is possible, though, that particular sound attributes are simply better suited to display particular data types (Barrass, 1997; B. N. Walker, 2000, 2002). A *categorical* sound dimension such as timbre (i.e., distinct musical instruments) might be well suitable to convey a categorical data distinction, such as different product lines (e.g., the L.A. office's hardware sales being played by a piano, and the software sales being played by a trumpet). That same dimension of timbre (or instrument) would be less suited to convey the variations in a continuous variable, such as sales volume. That is, we could have a sonification using the mappings {product:instrument, dollars:pitch}, but not {product:pitch, dollars:instrument}. Second, there may be a preferred polarity to the mapping. The intended listening population may agree (or at least, there may be a majority) that increasing frequency should be used to represent increasing sales volume. The issues surrounding such mapping choices have been discussed to some extent in the design literature, notably by Bertin (see also Barrass, 1997; 1981), however, this has not been a topic where the traditional psychoacoustics community has ventured (though, see more on this topic, below). Mapping is crucial, though, for auditory display designers in a practical setting where decisions must be made, so it is encouraging that new approaches have been made recently (Edworthy et al., 1991; B. N. Walker, 2000, 2002).

f) Rigid experimental methods and language

The psychoacoustics research field has provided many useful approaches and experimentation techniques. This, and the means to communicate the results of such experiments have been listed here already as beneficial contributions. However, it is a double-edged sword: The methodological rigor and experimental "purity" generally required in a psychophysical experiment may not be required, or even possible, in some auditory display research. It may be impossible, for example, to rigorously isolate a single independent variable in a sonification experiment involving multi-dimensional data represented by a highly multivariate sound. In some cases, such as assessing the effectiveness of a particular alarm, more qualitative methods, which can still be based on deep scientific foundations, may be appropriate. It is simply important for auditory display designers and researchers to cast their net wide in the search for useful and appropriate ways to evaluate their designs. Language presents further problems. It is sometimes limiting to feel compelled to describe sounds in the somewhat sterile terms of frequency and envelope, rather than with more descriptive terms like warmth, raspiness, or buzziness, especially if the participants in a study use those terms. Is seems somehow less scientific to write “more boingy”, even if that captures the phenomenological experience. This tension between science and, well, art, is not unique to auditory displays. Not infrequently, visual display designers struggle between the scientific world of terms like visual angle, information density, and contrast sensitivity functions, and the more vernacular terms like busy and cluttered. Auditory display designers need to be conversant in both forms of communication, and need to feel comfortable using both traditional and non-traditional methods of testing and evaluation. It may be less traditional, but it need not be less scientific.

6. What has recent research into higher-level phenomena offered to auditory display?

a) Gibsonian thinking in general

The first key contribution to auditory display from recent ecological approaches to perception and acoustics has been the general acceptance of the need to explore complex acoustic and perceptual phenomena and find new and appropriate means to do so. The possibilities have grown for conducting studies that do not hold all but one variable strictly constant, that are not necessarily performed in acoustically pure listening environments, and that involve more cognitive and even ill-defined tasks. This makes it feasible to study auditory display design and usage in realistic conditions, raising the validity and utility of the findings.

b) Complex, dynamic (ecological) sounds

Recent studies have begun to look at sounds that change dynamically within a trial. Walker and Kramer (1996b) took a fairly classic approach, employing reaction time and accuracy measures to study a fictional Crystal Factory sonification. However, the sounds were very dynamic, with one of several possible sound parameters varying continuously on each trial. Further, the experiment involved a fairly complex monitoring task, where the listener had to interpret the problem with the factory and make a corrective action. Thus, the participants were listening for “pressure” or “temperature” changes, and not “pitch” or “loudness”. This type of conceptual listening, and the inclusion of dynamic sounds, is not representative of traditional psychoacoustics studies but it certainly is relevant for auditory display tasks.

Walker and Ehrenstein (2000) also examined reaction times, accuracy, and stimulus-response (S-R) compatibility effects involving dynamic sounds that changed in pitch during a trial. The findings that pitch *change* can be considered a dimension (in addition to pitch, itself), and that S-R compatibility effects can arise from it, pointed out the need for designers of auditory displays to consider not only the perceptual issues resulting from their sounds, but also the response requirements.

Neuhoff and McBeath (1996) have also examined perception of dynamic sounds, notably those that change in pitch and loudness as if they are moving toward or away from the listener. The implications for train warning horns are clear; but the results more generally point out the over- and underestimation of pitch and loudness that occurs in these kinds of multi-dimensional sounds. For proper calibration, and maximum comprehension, auditory display designers need to consider these findings with dynamic sounds.

More realistic sounds are being used in otherwise “traditional” studies, as well. For example, Bonebright (2001) used multidimensional scaling techniques to study perception and categorization of environmental sounds. The results can help auditory display designers anticipate the conceptual groupings that will happen when such sounds are used in the interface, as with earcons, or auditory icons. Ballas (1994) and Gaver (1994) have also investigated the use of “real-world” sounds. These studies allow for a better understanding of the recognition of, memory for, and meanings attributed to common environmental sounds when used in an auditory display. In some cases, studying these reality-based sounds can also allow the designer to extract the salient features, and thereby use parameter-based synthesis models to create auditory icons that can have more in common (acoustically) with the objects they represent. Thus, the study of realistic sounds in some cases allows auditory designers to use those with a better level of understanding, and in some cases allows designers to create artificial sounds that are more efficient, but equally compelling and useful.

c) Acoustic ecologies for overall design

As the use of sound in human-machine interfaces increases, it is even more important to know how individual sounds will blend in or stand out from the growing acoustic crowd. However, it is also important to take into consideration the whole sound environment, and design a total auditory display environment that supports the tasks, as well as the persons in that environment. These design issues have recently been considered, though perhaps still less rigorously than they might, by the auditory perception community. For example, Walker and Kramer (1996a) point out that the nature of the human-system interaction ("concert" versus "conversation" modes) can inform the auditory display needs. They recommend that auditory (and visual) displays be designed to be adaptive to the interaction style. Some auditory display designers are now considering these issues in their designs, under the aegis of "contextual computing" (e.g., Nagel, Kidd, O'Connell, Dey, & Abowd, 2001), and are developing displays that adapt to the location, activities, and preferences of the intended listener.

d) Metaphors and conceptual mappings

Traditional psychoacoustics has long been concerned with the issue of how listeners' perceptions of a sound (e.g., pitch) compare to the sound's physical parameter (e.g., frequency). Psychophysical scaling functions have been determined for all kinds of sound dimensions (e.g., Stevens, 1975). As auditory displays fill out more of the conceptual and symbolic end of the spectrum, it is crucial to know how listeners are interpreting the sounds, not just how they perceive them. Walker has been using magnitude estimation to determine psychophysical scaling functions between sound dimensions (frequency, tempo, brightness), and *conceptual* data dimensions, like pressure, temperature, and number of dollars (B. N. Walker, 2000, 2002). It turns out that the best mappings depend not only on what the sound is, but also what the data dimension is. For example, if frequency is used to represent temperature, the preferred polarity is positive. That is, increasing frequency represents increasing temperature. However, increasing frequency best represents decreasing size (B. N. Walker, 2000, 2002). In addition to the findings regarding polarity, this research shows that the slope of the magnitude estimation graph can be different for each data dimension. For example, a given change in frequency may best indicate a certain percentage change in temperature, but a different change in pressure or number of dollars. Both the polarity preferences and scaling factors are now being used to create auditory displays that conform better to population expectancies (Smith & Walker, 2002; B. N. Walker, 2002).

One additional contribution from this line of research has been the finding that different listener populations can interpret auditory displays differently. Most notably, Walker and Lane (2001) found that visually impaired listeners used different mental models to interpret the sounds they heard than did sighted listeners. For example, sighted undergraduates preferred the positive polarity mapping for frequency:number of dollars. That is, a higher frequency meant a greater number of dollars. For the blind listeners, increasing frequency unanimously meant *fewer* dollars. The study points out that different listening experience can affect how we interpret the sounds we hear. These results have practical utility for auditory display designers in that the designers can use different mappings for blind listeners. However, the research is probably just as important for its "philosophical" contribution, pointing out that designers really do need to "know their users" at a more cognitive level than had been addressed in traditional studies that generally diminished the importance of learning and experience on interpretation of sounds.

e) Alternative experimental methods

The general trend has been for researchers interested in auditory displays to use somewhat traditional methods (psychophysical scaling, magnitude estimation, multidimensional scaling), but employ more dynamic sounds, non-traditional listening environments, or more cognitive tasks. However, there have been some new experimental approaches, such as the "ventriloquism effect" for multimodal displays (A. Walker & Brewster, 2001), "audio descriptive analysis" for spatial sound displays (Zacharov & Koivuniemi, 2001), and the use of auditory displays as part of the experimental task (e.g., a crystal factory, B. N. Walker & Kramer, 1996b). Traditional experimental approaches would not have been very well suited to uncovering mental models, and previous applications might not have even cared much about that. Auditory displays require a new game plan.

7. Where are we now and where do we need to go vis a vis auditory display?

Research in auditory perception has progressed from the study of individual auditory dimensions, such as pitch, tempo, loudness, and localization, to the study of more complex phenomena, such as auditory streaming, dynamic sound perception, auditory attention, and multimodal displays. The main research areas that will continue to drive sonification research forward include (1) understanding dynamic sound perception, (2) investigating auditory streaming, (3) cognitive issues in the relationship between conceptual (data) and auditory (display) features, (4) determining how meaning is attributed to sounds, based on the sounds, the listener, and the task requirements, and (5) understanding the design and use of multimodal sonification.

Other pertinent research includes investigation of the normal variation in perceptual and cognitive abilities and strategies in the human population, differences in cognitive representations of auditory displays for selected populations, such as sighted and blind individuals or different cultural groups, and the role of learning and familiarity in display efficacy. Study of these issues is necessary for the formulation of design guidelines for constructing efficient multivariate data displays using sounds.

Substantial questions remain in the area of multimodal perception. It is clear from existing research that our senses differ in their underlying abilities, but further research is necessary to identify what data features are perceptually most salient to each sense and to determine how to use this knowledge in designing effective displays. Multimodal interactions (e.g., between visual and auditory displays) are poorly understood, yet they critically affect most sonification applications. When does redundant presentation (that is, presenting information in more than one modality) improve the ability to extract data (i.e., cross-modal synergy)? When does information presented in one modality interfere with the perception of information in another modality (i.e., cross-modal interference)? How can the total amount of information perceived across all modalities be maximized?

Synergies between Auditory Display and Ecological Psychoacoustics Research

Many of the questions presented here regarding Ecological Psychoacoustics, and other questions as yet unformulated, may be best approached via the field of Auditory Display. That is, abstract questions regarding auditory perception, such as “How well can a person discriminate these three auditory variables?” may be more effectively approached by practical questions such as, “Can users of this three-variable display perform a task which involves discerning the relative values of each of the three data dimensions being represented?” That is, the perceptual questions are addressed via a sonification task, thereby assuring that the results will be clearly relevant to auditory displays. At the same time, the perceptual questions will yield findings that also have broad theoretical relevance.

It also seems likely that auditory display research will present new questions and perspectives to ecological psychoacoustics researchers. Display designs may present novel cognitive tasks, such as cross-modal pattern detection or discerning trends in highly multivariate displays. Fortunately, the displays presenting these perceptual quandaries will often also provide the experimental tools for their exploration. Thus a synergy is created wherein auditory display research pushes the boundaries of ecological psychoacoustics, and ecological psychoacoustics informs and validates auditory display research. Both fields are thus broadened, refined, and brought to greater practical application.

References

Astheimer, P. (1993). Realtime sonification to enhance the human-computer interaction in virtual worlds. *Proceedings of the Fourth Eurographics Workshop on Visualization in Scientific Computing*. Abingdon, England.

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology, 19*(2), 250-267.

Ballas, J. A. (1994). Delivery of information through sound. In G. Kramer (Ed.), *Auditory display: sonification, audification, and auditory interfaces* (pp. 79-94). Reading, MA: Addison-Wesley.

Barrass, S. (1997). *Auditory information design*. Unpublished Dissertation, Australian National University.

Bertin, J. (1981). *Graphics and graphic information processing*. Berlin: Walter de Gruyter.

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. *Human-Computer Interaction, 4*, 11-44.

Bly, S. (1982). *Sound and Computer Information Presentation*. Unpublished Ph.D. Dissertation, University of California, Davis.

Bonebright, T. (2001). Perceptual structure of everyday sounds: A multidimensional scaling approach. *Proceedings of the 7th International Conference on Auditory Display* (pp. 73-78), Espoo, Finland.

Bregman, A. S. (1990). *Auditory scene analysis : the perceptual organization of sound.* Cambridge, MA: MIT Press.

Brungart, D. S., Ericson, M. A., & Simpson, B. D. (2002). Design considerations for improving the effectiveness of multitalker speech displays. *Proceedings of the 8th International Conference on Auditory Display* (pp. 424-430), Kyoto, Japan.

Carlile, S., Leong, P., & Hyams, S. (1997). The nature and distribution of errors in sound localization by human listeners. *Hearing Research, 114*, 79-196.

Dombois, F. (2001). Using audification in planetary seismology. *Proceedings of the 7th International Conference on Auditory Display* (pp. 227-230), Espoo, Finland.

Dombois, F. (2002). Auditory seismology - On free oscillations, focal mechanisms, explosions, and synthetic seismograms. *Proceedings of the 8th International Conference on Auditory Display* (pp. 27-30), Kyoto, Japan.

Earl, C., & Leventhal, J. A. (1999). Survey of Windows screen reader users: Recent improvements in accessibility. *Journal of Visual Impairment and Blindness, 93*(3).

Edworthy, J., Loxley, S., & Dennis, I. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human Factors, 33*, 205-232.

Egan, J., Carterette, E., & Thwing, E. (1954). Some factors affecting multichannel listening. *Journal of the Acoustical Society of America, 26*, 774-782.

Fitch, W. T., & Kramer, G. (1994). Sonifying the body electric: Superiority of an auditory over a visual display in a complex, multi-variate system. In G. Kramer (Ed.), *Auditory display: sonification, audification, and auditory interfaces* (pp. 307-326). Reading, MA: Addison-Wesley.

Fletcher, H. (1940). Auditory patterns. *Reviews of Modern Physics, 12*, 47-65.

Fletcher, H., & Munson, W. A. (1933). Loudness: Its definition, measurement, and calculation. *Journal of the Acoustical Society of America, 5*, 82-88.

Flowers, J. H., & Grafel, D. C. (2002). Perception of sonified daily weather records. *Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society*. Baltimore, Maryland.

Garner, W. R. (1974). *Processing of information and structure*. Hillsdale, NJ: Erlbaum.

Gaver, W. W. (1989). The SonicFinder: An interface that uses auditory icons. *Human-Computer Interaction, 4*(1), 67-94.

Gaver, W. W. (1994). Using and creating auditory icons. In G. Kramer (Ed.), *Auditory display: sonification, audification, and auditory interfaces* (pp. 417-446). Reading, MA: Addison-Wesley.

Gaver, W. W., Smith, R. B., & O'Shea, T. (1991). Effective sounds in complex systems: the ARKola simulation. *Proceedings of the CHI'91*. New York.

Gittins, D. (1986). Icon-based human-computer interaction. *International Journal of Man-Machine Studies, 24*, 519-543.

Guillaume, A., Drake, C., Rivenez, M., Pellieux, L., & Chastres, V. (2002). Perception of urgency and alarm design. *Proceedings of the 8th International Conference on Auditory Display* (pp. 357-361), Kyoto, Japan.

Hartmann, W. M. (1997). *Sounds, signals, and sensation: Modern acoustics and signal processing*. New York: Springer Verlag.

Hayward, C. (1994). Listening to the earth sing. In G. Kramer (Ed.), *Auditory display: sonification, audification, and auditory interfaces* (pp. 369-404). Reading, MA: Addison-Wesley.

Hellier, E. J., Edworthy, J., & Dennis, I. (1993). Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. *Human Factors, 35*, 693-706.

Hemenway, K. (1982). Psychological issues in the use of icons in command menus. *Proceedings of the CHI'82 Conference on Human Factors in Computer Systems* (pp. 21-24), New York.

Howard, J. H., & Ballas, J. A. (1982). Acquisition of acoustic pattern categories by exemplar observation. *Organizational Behavior and Human Decision Processes, 30*(2), 157-173.

ISO (1987). *Acoustics -- Normal equal-loudness level contours* (No. ISO 226:1987). Geneva: International Organization for Standardization.

Kolers, P. (1969). Some formal characteristics of pictograms. *American Scientist, 57*(3), 348-363.

Kramer, G. (1993). Sonification of Financial Data: An Overview of Spreadsheet and Database Sonification. *Proceedings of the Virtual Reality Systems '93* (pp. New York.

Kramer, G. (1994a). An introduction to auditory display. In G. Kramer (Ed.), *Auditory display: Sonification, audification, and auditory interfaces* (pp. 1-78). Reading, MA: Addison Wesley.

Kramer, G. (1994b). Some organizing principles for representing data with sound. In G. Kramer (Ed.), *Auditory display: Sonification, audification, and auditory interfaces* (pp. 185-222). Reading, MA: Addison Wesley.

Kramer, G., & Ellison, S. (1991). Audification: The use of sound to display multivariate data. *Proceedings of the International Computer Music Conference* (pp. 214-221), San Francisco, CA.

Kramer, G., Walker, B. N., Bonebright, T., Cook, P., Flowers, J., Miner, N., et al. (1999). *The Sonification Report: Status of the Field and Research Agenda. Report prepared for the National Science Foundation by members of the International Community for Auditory Display*. Santa Fe, NM: International Community for Auditory Display (ICAD).

Licklider, J. C. R. (1951). Basic correlates of the auditory stimulus. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 985-1039). New York: Wiley & Sons.

Licklider, J. C. R. (1956). Auditory frequency analysis. In C. Cherry (Ed.), *Information theory* (pp. 253-268). New York: Academic Press.

Martins, A. C. G., & Rangayyan, R. M. (1997). Experimental evaluation of auditory display and sonification of textured images. *Proceedings of the Fourth International Conference on Auditory Display (ICAD97)* (pp. 129-134), Palo Alto, CA.

McAdams, S., & Bigand, E. (Eds.). (1993). *Thinking in sound : the cognitive psychology of human audition*. Oxford England ; New York: Clarendon Press.

McKinley, R. L., & Ericson, M. A. (1992). Minimum audible angles for synthesized localization cues presented over headphones. *Journal of the Acoustical Society of America, 92*(4), 2297.

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology, 42*, 135-159.

Mills, A. W. (1958). On the minimum audible angle. *Journal of the Acoustical Society of America, 30*, 237-246.

Moore, B. C. J. (1989). *An introduction to the psychology of hearing*. London: Academic Press.

Mowbry, G. H., & Gebhard, J. W. (1961). Man's senses as informational channels. In H. W. Sinaiko (Ed.), *Human factors in the design and use of control systems* (pp. 115-149). New York: Dover.

Nagel, K., Kidd, C. C., O'Connell, T., Dey, A., & Abowd, G. D. (2001). The Family Intercom: Developing a Context-Aware Audio Communication System. *Proceedings of the Ubicomp 2001* (pp. 176-183).

Neuhoff, J., & McBeath, M. K. (1996). The Doppler Illusion: The influence of dynamic intensity change on perceived pitch. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 970-985.

OSHA (1981). *OSHA Standard 1910.95: Occupational noise exposure*: Occupational Safety and Health Administration.

Ossner, J. (1990). Transnational Symbols: The rule of pictograms and models in the learning process. In J. Nielsen (Ed.), *Designing user interfaces for international use* (pp. 11-38). Amsterdam: Elsevier.

Patterson, R. D. (1982). *Guidelines for auditory warning systems on Civil Aircraft.* London: Civil Aviation Authority.

Pereverzev, S. V., Loshak, A., Backhaus, S., Davis, J. C., & Packard, R. E. (1997). Quantum oscillations between two weakly coupled reservoirs of superfluid 3He. *Nature, 388*, 449-451.

Pollack, I., & Ficks, L. (1954). Information of elementary multidimensional auditory displays. *Journal of the Acoustical Society of America, 26*, 155-158.

Proctor, R. W., & Reeve, T. G. (Eds.). (1990). *Stimulus-response compatibility: An integrated perspective*. Amsterdam: North Holland.

Resnick, S. B., & Feth, L. L. (1975). Discriminability of time-reversed click pairs: intensity effects. *Journal of the Acoustical Society of America, 57*, 1493-1499.

Robinson, D. W., & Dadson, R. S. (1956). A re-determination of the equal-loudness relations for pure tones. *British Journal of Applied Physics, 7*, 166-181.

Salvendy, G. (1997). *Handbook of human factors and ergonomics* (2nd ed.). New York: Wiley.

Sanders, M. S., & McCormick, E. J. (1993). *Human factors in engineering and design* (7th ed.). New York: McGraw-Hill.

Scarf, F. L. (1979). Possible Traversals of Jupiter's Distant Magnetic Tail by Voyager and by Saturn. *Journal of Geophysical Research, 84*, 4422.

Schiffman, H. R., Reynolds, M. L., & Young, F. L. (1981). *Introduction to multidimensional scaling*. New York: Academic Press.

Schouten, J. F. (1940). The residue and the mechanism of hearing. *Proc. K. Ned. Akad. Wet., 43*, 991-999.

Shneiderman, B. (1998). *Designing the user interface : strategies for effective human-computer-interaction* (3rd ed.). Reading, MA: Addison Wesley Longman.

Smith, D. R., & Walker, B. N. (2002, 02-05 July). Tick-marks, axes, and labels: The effects of adding context to auditory graphs. *Proceedings of the 8th International Conference on Auditory Display* (pp. 362-367), Kyoto, Japan.

Sorkin, R. D. (1987). Design of auditory and tactile displays. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 549-576). New York: Wiley & Sons.

Speeth, S. D. (1961). Seismometer sounds. *Journal of the Acoustical Society of America, 33*, 909-916.

Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects*. New York: Wiley.

Treisman, A. (1964). The effects of irrelevant material on the efficiency of selective listening. *American Journal of Psychology, 77*, 533-546.

Uzilevsky, G., & Andreev, V. (1993). Iconic signs and languages in user interface development. In L. J. Bass, J. Gomostaev & C. Unger (Eds.), *Human-Computer Interaction. Selected Papers from Third International Conference, EWHCI'93* (pp. 115-124). Heidelberg: Springer-Verlag.

Valenzuela, M. L., Sansalone, M. J., Krumhansl, C. L., & Streett, W. B. (1997). Use of sound for the interpretation of impact-echo signals. *Proceedings of the Fourth International Conference on Auditory Display (ICAD97)* (pp. 47-56), Palo Alto, CA.

Walker, A., & Brewster, S. (2001). "Sitting too close to the screen can be bad for your ears": A study of audio-visual location discrepancy detection under different visual projections. *Proceedings of the 7th International Conference on Auditory Display* (pp. 86-89), Espoo, Finland.

Walker, B. N. (2000). *Magnitude estimation of conceptual data dimensions for use in sonification*. Unpublished Ph.D. Dissertation, Rice University, Houston, TX.

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

Walker, B. N., & Ehrenstein, A. (2000). Pitch and pitch change interact in auditory displays. *Journal of Experimental Psychology: Applied, 6*, 15-30.

Walker, B. N., & Kramer, G. (1996a). Human factors and the acoustic ecology: Considerations for multimedia audio design. *Proceedings of the Audio Engineering Society 101st Convention*. Los Angeles.

Walker, B. N., & Kramer, G. (1996b). Mappings and metaphors in auditory displays: an experimental assessment. *Proceedings of the 3rd International Conference on Auditory Display (ICAD96)* (pp. 71-74), Palo Alto, CA.

Walker, B. N., & Lane, D. M. (2001). Psychophysical scaling of sonification mappings: A comparision of visually impaired and sighted listeners. *Proceedings of the 7th International Conference on Auditory Display* (pp. 90-94), Espoo, Finland.

Wenzel, E. M. (1994). Spatial sound and sonification. In G. Kramer (Ed.), *Auditory display: sonification, audification, and auditory interfaces* (pp. 127-150). Reading, MA: Addison-Wesley.

Zacharov, N., & Koivuniemi, K. (2001). Audio descriptive analysis and mapping of spatial sound displays. *Proceedings of the 7th International Conference on Auditory Display* (pp. 95-104), Espoo, Finland.

Zwicker, E., & Scharf, B. (1965). A model of loudness summation. *Psychological Review, 72*, 3-26.