

## 2 Theory of Sonification

Bruce N. Walker and Michael A. Nees

Georgia Institute of Technology, Atlanta, GA, USA  
bruce.walker@psych.gatech.edu; mnees@gatech.edu

### 2.1 Chapter Overview

Auditory displays can be broadly defined as any display that uses sound to communicate information. Sonifications are a subtype of auditory displays that use nonspeech audio to represent information. Kramer et al. (1999) further elaborated that “sonification is the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation.” Sonification, then, seeks to translate relationships in data into sound(s) that exploit the auditory perceptual abilities of human beings such that the data relationships are comprehensible.

Sonification is a truly interdisciplinary approach to information display, and, as Kramer (1994) pointed out, a complete understanding of the field would require many lifetimes of expertise across many domains of knowledge. The theoretical underpinnings of research and design for sonification comes from such diverse fields as audio engineering, audiology, computer science, informatics, linguistics, mathematics, music, psychology, and telecommunications, to name but a few, and are as yet not characterized by a single grand or unifying set of principles or rules (see Edworthy, 1998). Rather, the guiding theory (or theories) of sonification in practice can be best characterized as an amalgam of important insights drawn from the convergence of these many diverse fields.

The 1999 collaborative *Sonification Report* (Kramer et al., 1999) identified four issues that should be addressed in a theoretical description of sonification. These included: (1) taxonomic descriptions of sonification techniques based on psychological principles or display applications; (2) descriptions of the types of data and user tasks amenable to sonification; (3) a treatment of the mapping of data to acoustic signals; and (4) a discussion of the factors limiting the use of sonification. By addressing the current status of these four topics, the current chapter seeks to provide a broad introduction to sonification, as well as an account of the guiding theoretical considerations for sonification researchers and designers. Dozens of active contributors from multiple disciplines have collectively established a solid base of knowledge in sonification research and design. This knowledge base reflects the multifaceted

nature of sonification, and this collaborative, multidisciplinary approach to the field has allowed us today to lay down an overview of the principles employed in sonification research and design. We attempt to draw upon the insights of relevant domains of research, and where necessary, offer areas where future researchers could answer unresolved questions or make fruitful clarifications or qualifications to the current state of the field. In many cases, we will point the interested reader to another more detailed chapter in this book, or to other external sources for more extensive coverage.

## 2.2 Sonification and Auditory Displays

Sonifications are a relatively recent subset of auditory displays. As in any information system (see Figure 2.1), an auditory display offers a relay between the information source and the information receiver (see Kramer, 1994; Shannon, 1998/1949). In the case of an auditory display, the data of interest are conveyed to the human listener through sound.

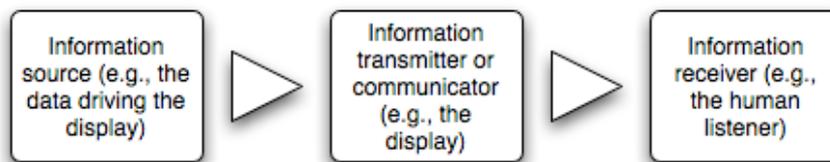


Figure 2.1: General description of a communication system

Although investigations of audio as an information display date back over 50 years (see Frysinger, 2005), digital computing technology has more recently meant that auditory displays of information have become ubiquitous. Edworthy (1998) argued that the advent of auditory displays and audio interfaces was practically inevitable given the ease and cost efficiency with which computers can now produce sound. Devices ranging from cars to computers to cell phones to microwaves pervade our environments, and all of these devices now use intentional sound<sup>1</sup> to deliver messages to the user.

The rationales and motivations for displaying information using sound (rather than a visual presentation, etc.) have been discussed at length elsewhere. Briefly, though, auditory displays exploit the superior ability of the human

---

<sup>1</sup> *Intentional* sounds are purposely engineered to perform as an information display (see Walker & Kramer, 1996), and stand in contrast to *incidental* sounds, which are non-engineered sounds that occur as a consequence of the normal operation of a system (e.g., a car engine running). Incidental sounds may be quite informative (e.g., the sound of wind rushing past can indicate a car's speed), though this characteristic of incidental sounds is serendipitous rather than designed. The current chapter is confined to a discussion of intentional sounds.

auditory system to recognize temporal changes and patterns (Bregman, 1990; Flowers, Buhman, & Turnage, 1997; Flowers & Hauer, 1995; Garner & Gottwald, 1968; Kramer et al., 1999; McAdams & Bigand, 1993; Moore, 1997). As a result, auditory displays may be the most appropriate modality when the information being displayed has complex patterns, changes in time, includes warnings, or calls for immediate action. Second, in practical work environments the operator is often unable to *look* at, or unable to *see*, a visual display. The visual system might be busy with another task (Fitch & Kramer, 1994; Wickens & Liu, 1988), or the perceiver might be visually impaired, either physically or as a result of environmental factors such as smoke or line of sight (Fitch & Kramer, 1994; Kramer et al., 1999; Walker, 2002; Walker & Kramer, 2004; Wickens, Gordon, & Liu, 1998), or the visual system may be overtaxed with information (see Brewster, 1997; M. L. Brown, Newsome, & Glinert, 1989). Third, auditory and voice modalities have been shown to be most compatible when systems require the processing or input of verbal-categorical information (Salvendy, 1997; Wickens & Liu, 1988; Wickens, Sandry, & Vidulich, 1983). Other features of auditory perception that suggest sound as an effective data representation technique include our ability to monitor and process multiple auditory data sets (parallel listening) (Fitch & Kramer, 1994), and our ability for rapid auditory detection, especially in high stress environments (Kramer et al., 1999; Moore, 1997). Finally, with mobile devices becoming increasingly smaller in size, sound may be a compelling display mode as visual displays decrease in size (Brewster & Murray, 2000). For a more complete discussion of the benefits of (and potential problems with) auditory displays, see Kramer (1994; Kramer et al., 1999), Sanders and McCormick (1993), Johannsen (2004), and Stokes (1990).

### **2.3 Towards a Taxonomy of Auditory Display & Sonification**

A taxonomic description of auditory displays in general, and sonifications in particular, could be organized in any number of ways. Categories often emerge from either the function of the display or the technique of sonification, and either could serve as the logical foundation for a taxonomy. In this chapter we offer a discussion of ways of classifying auditory displays and sonifications according to both function and technique, although, as our discussion will elaborate, they are very much inter-related.

Sonifications are clearly a subset of auditory display, but it is not clear, in the end, where the exact boundaries should be drawn. Categorical definitions within the sonification field tend to be loosely enumerated and are somewhat flexible. For example, auditory representations of box-and-whisker plots, diagrammatic information, and equal-interval time series data have all been called sonification, and, in particular, “auditory graphs,” but all of these displays are clearly quite different from each other in both form and function. Ultimately, the name assigned to a sonification is much less important than its ability to

communicate the intended information. Thus, the taxonomic description that follows is intended to parallel conventional naming schemes found in the literature, but these descriptions should not be taken to imply that clear cut boundaries and distinctions are always possible to draw, nor are they crucial to the creation of a successful display.

### **2.3.1 Functions of sonification**

Given that sound has some inherent properties that should prove beneficial as a medium for information display, we can begin by considering some of the functions that auditory displays might perform. Buxton (1989) and others (e.g., Edworthy, 1998; Kramer, 1994; Walker & Kramer, 2004) have described the function of auditory displays in terms of three broad categories: (1) alarms, alerts, and warnings; (2) status, process, and monitoring messages; and (3) data exploration. To this we would add: (4) art and entertainment.

#### *2.3.1.1 Alerting functions*

*Alerts and notifications* refer to sounds used to indicate that something has, or is about to occur, or that the listener should immediately attend to something in the environment (see Buxton, 1989; Sanders & McCormick, 1993; Sorkin, 1987). Alerts and notifications tend to be simple and particularly overt. The message conveyed is information-poor. For example, a beep is often used to indicate that the cooking time on a microwave oven has expired. There is generally little information as to the details of the event—the microwave beep merely indicates that the time has expired, not necessarily that the food is fully cooked. Another commonly heard alert is a doorbell—the basic ring does not indicate who is at the door, or why.

*Alarms and warnings* are alert or notification sounds that are intended to convey the occurrence of a constrained class of events, usually adverse, that carry particular urgency in that they require immediate response or attention (see Haas & Edworthy, 2006). Warning signals presented in the auditory modality automatically capture spatial attention better than visual warning signals (Spence & Driver, 1997). A well-chosen alarm or warning should, by definition, carry slightly more information than a simple alert (i.e., the user knows that an alarm indicates an adverse event that requires an immediate action); however, the specificity of the information about the adverse event generally remains limited. Fire alarms, for example, signal an adverse event (a fire) that requires immediate action (evacuation), but the alarm does not carry information about the location of the fire or its severity.

More complex (and modern) kinds of alarms attempt to encode more information into the auditory signal. Examples range from families of categorical warning sounds in healthcare situations (e.g., Sanderson, in press) to helicopter telemetry and avionics data being used to modify a given warning sound (e.g., “trendsons”, Edworthy, Hellier, Aldrich, & Loxley, 2004). These sounds,

discussed at length by Edworthy and Hellier (2006), blur the line between alarms and status indicators, discussed next.

#### *2.3.1.2 Status and progress indicating functions*

Although in some cases sound performs a basic alerting function, other scenarios require a display that offers more detail about the information being represented with sound. The current or ongoing status of a system or process often needs to be presented to the human listener, and auditory displays have been applied as dynamic *status and progress indicators*. In these instances, sound takes advantage of “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 uses ranging from monitoring models of factory process states (see Gaver, Smith, & O'Shea, 1991; Walker & Kramer, 2005), to patient data in an anesthesiologist's workstation (Fitch & Kramer, 1994), blood pressure in a hospital environment (M. Watson, 2006), and telephone hold time (Kortum, Peres, Knott, & Bushey, 2005).

#### *2.3.1.3 Data exploration functions*

The third functional class of auditory displays are those designed to permit *data exploration*. These are what is generally meant by the term “sonification”, and are usually intended to encode and convey information about an entire data set or relevant aspects of the data set. Sonifications designed for data exploration differ from status or process indicators in that they use sound to offer a more holistic portrait of the data in the system rather than condensing information to capture a momentary state such as with alerts and process indicators. *Auditory graphs* (for representative work, see L. M. Brown & Brewster, 2003; Flowers & Hauer, 1992, 1993, 1995; Smith & Walker, 2005) and interactive sonifications (see Chapter 13 in this volume and Hermann & Hunt, 2005) are typical exemplars of sonifications designed for data exploration purposes.

#### *2.3.1.4 Art and entertainment*

As the sound-producing capabilities of computing systems have evolved, so too has the field of computer music. In addition to yielding warnings and sonifications, events and data sets can be used as the basis for musical compositions. Often the resulting performances include a combination of the types of sounds discussed to this point, in addition to more traditional musical elements. While the composers often attempt to convey something to the listener through these sonifications, it is not for the pure purpose of information delivery. Recent examples of sonification compositions have ranged from sonifications of human electroencephalogram (EEG) data (“Listening to the mind listening: Concert of sonifications at the Sydney Opera House”, 2004), to global economic

and health data ("Global music - The world by ear", 2006), among others. Quinn (2001, 2003) has used data sonifications to drive ambitious musical works, and he has published entire albums of his compositions.

### **2.3.2 Sonification techniques and approaches**

de Campo (2006) offered a sonification design map that featured three broad categorizations of sonification approaches: (1) event-based; (2) model-based; and (3) continuous. Again, the definitional boundaries to taxonomic descriptions of sonifications are indistinct and often overlapping. We provide a brief overview of approaches and techniques employed in sonification below; for a more detailed treatment, see later chapters in this volume.

#### *2.3.2.1 Modes of interaction*

A prerequisite to a discussion of sonification approaches is a basic understanding of the nature of the interaction that may be available to a user of an auditory display. Interactivity can be considered as a dimension along which different displays can be classified, ranging from completely non-interactive to completely user-initiated. For example, in some instances the listener may passively take in a display without being given the option to actively manipulate the display (by controlling the speed of presentation, pausing, fast-forwarding, or rewinding the presentation, etc.). The display is simply triggered and plays in its entirety while the user listens. Sonifications at this non-interactive end of the dimension have been called "concert mode" (Walker & Kramer, 1996) or "tour based" (Franklin & Roberts, 2004). Alternatively, the listener may be able to actively control the presentation of the sonification. In some instances, the user might be actively choosing and changing presentation parameters of the display (see L. M. Brown, Brewster, & Riedel, 2002). In other cases, user input and interaction may be the required catalyst that drives the presentation of sounds (see Hermann & Hunt, 2005). Sonifications more toward this interactive end of the spectrum have been called "conversation mode" (Walker & Kramer, 1996) or "query based" (Franklin & Roberts, 2004) sonification, and include "interactive sonification" (see Chapter 13 in this volume and Hermann & Hunt, 2005). Walker has pointed out that for most sonifications to be useful (and certainly those intended to support learning and discovery), there needs to be at least some kind of interaction capability, even if it is just the ability to pause or replay a particular part of the sound (e.g., Walker & Cothran, 2003; Walker & Lowey, 2004).

#### *2.3.2.2 Event-based sonification*

Event-based approaches to sonification describe those displays where the data are such that parameter mapping can be employed (de Campo, 2006; Hermann & Hunt, 2005). *Parameter mapping* represents changes in some data

dimension with changes in an acoustic dimension to produce a sonification (Hermann & Hunt, 2005). By definition, sonification represents changes in data with changes in one or more sound attributes (Kramer et al., 1999). Manipulatable perceptual dimensions of sound, therefore, must be mapped to correspond to changes in data. Sound, however, has a multitude of changeable dimensions (see Kramer, 1994; Levitin, 1999) that allow for a large design space when mapping data to audio. In order for parameter mapping to be used in a sonification, the dimensionality of the data must be constrained such that a perceivable display is feasible, thus parameter mapping tends to result in a more low dimension display than the model-based approaches discussed below. Event-based approaches to sonification have typically employed a somewhat passive mode of interaction. Indeed, some event-based sonifications (e.g., alerts and notifications, etc.) are designed to be brief and would offer little opportunity for user interaction. Other event-based approaches that employ parameter mapping for purposes of data exploration (e.g., auditory graphs) could likely benefit from adopting some combination of passive listening and active listener interaction.

#### *2.3.2.3 Model-based sonification*

Model-based approaches to sonification differ from event-based approaches in that instead of mapping data parameters to sound parameters, the display designer builds a virtual model with which the listener interacts such that the model's "properties are informed by the data" (de Campo, 2006, p. 2). A model constitutes a virtual object with which the user can interact, and the user's input drives the sonification such that the model is "a dynamic system capable of a dynamic behavior that can be perceived as sound" (Bovermann, Hermann, & Ritter, 2006, p. 78). The user comes to understand the structure of the data based on the acoustic responses of the model during interactive probing of the virtual object (Hermann & Hunt, 2005). Model-based approaches rely heavily upon the active manipulation of the sonification by the user and tend to involve high data dimensionality.

#### *2.3.2.4 Continuous sonification*

Continuous sonification may be possible when data are time series and sampled at a rate such that a quasi-analog signal can be directly translated into sound (de Campo, 2006). *Audification* is the most prototypical method of continuous sonification, whereby waveforms of periodic data are directly translated into sound (Kramer, 1994). For example, seismic data have been audified in order to facilitate the categorization of seismic events with accuracies of over 90% (see Dombois, 2002; Speeth, 1961). This approach may require that the waveforms be frequency- or time-shifted into the range of audible waveforms for humans.

*2.3.2.5 The convergence of taxonomies of function and technique*

Although accounts to date have generally classified sonifications in terms of function or technique, the categorical boundaries of functions and techniques are vague. Furthermore, the function of the display in a system may constrain the sonification technique, and the choice of technique may limit the functions a display can perform. Event-based approaches are the only approach used for alerts, notifications, alarms, and even status and process monitors, as these functions are all event-based. Data exploration may employ event-based approaches, model-based sonification, or continuous sonification depending upon the specific task of the user.

## **2.4 Data Properties and Task Dependency**

The nature of the data to be presented and the task of the human listener are important factors for a system that employs sonification for information display. The display designer must consider, among other things: what the user needs to accomplish (i.e., the task(s)); what parts of the information source (i.e., the data) are relevant to the user's task; how much information the user needs to accomplish the task; what kind of display to deploy (simple alert, status indicator, or full sonification, for example); and how to manipulate the data (e.g., filtering, transforming, or data reduction).

These issues come together to present major challenges in sonification design, since the nature of the data and the task will necessarily constrain the data-to-display mapping design space. Part of this is perceptual or "bottom up", in that some dimensions of sound are perceived as categorical (e.g., timbre), whereas other attributes of sound are perceived along a perceptual continuum (e.g., frequency, intensity). Part of the challenge comes from the more cognitive or conceptual "top down" components of sonification usage. For example, Walker (2002) has shown that conceptual dimensions (like size, temperature, price, etc.) influence how a listener will interpret and scale the data-to-display relationship.

### **2.4.1 Data types**

Information can be broadly classified as quantitative (numerical) or qualitative (verbal), and the design of an auditory display to accommodate quantitative data may be quite different from the design of a display that presents qualitative information. Data can also be described in terms of the scale upon which measurements were made. Nominal data classify or categorize; no meaning beyond group membership is attached to the magnitude of numerical values for nominal data. Ordinal data take on a meaningful order with regards to some quantity, but the distance between points on ordinal scales may vary. Interval and ratio scales have the characteristic of both meaningful order and meaningful distances between points on the scale (see S.S. Stevens, 1946). Data

can also be discussed in terms of its existence as discrete pieces of information (e.g., events or samples) versus a continuous flow of information.

Barrass (1997, 2005) is one of the few researchers to consider the role of different types of data in auditory display and make suggestions about how information type can influence mappings. As one example, nominal/categorical data types (e.g., different cities) should be represented by categorically changing acoustic variables, such as timbre. Interval data may be represented by more continuous acoustic variables, such as pitch or loudness (but see S. S. Stevens, 1975; Walker, in press, for more discussion on this issue).

Nevertheless, there remains a paucity of research aimed at studying the factors within a data set that can affect perception or comprehension. For example, data that are generally slow-changing, with relatively few inflection points (e.g., rainfall or temperature) might be best represented with a different type of display than data that are rapidly-changing with many direction changes (e.g., EEG or stock market activity). Presumably, though, research will show that data set characteristics such as density and volatility will affect the best choices of mapping from data to display. This is beginning to be evident in the work of Hermann, Dombois, and others who are using very large and rapidly changing data sets, and are finding that audification and model-based sonification are more suited to handle them. Even with sophisticated sonification methods, data sets often need to be pre-processed, reduced in dimensionality, or sampled to decrease volatility before a suitable sonification can be created. On the other hand, smaller and simpler data sets such as might be found in a high-school science class may be suitable for direct creation of auditory graphs and auditory histograms.

#### **2.4.2 Task types**

Task refers to the functions that are performed by the human listener within a system like that depicted in Figure 2.1. Although the most general description of the listener's role involves simply receiving the information presented in a sonification, the person's goals and the functions allocated to the human being in the system will likely require further action by the user upon receiving the information. Furthermore, the auditory display may exist within a larger acoustic context in which attending to the sound display is only one of many functions concurrently performed by the listener. Effective sonification, then, requires an understanding of the listener's function and goals within a system. What does the human listener need to accomplish? Given that sound represents an appropriate means of information display, how can sonification best help the listener successfully perform her or his role in the system? Task, therefore, is a crucial consideration for the success or failure of a sonification, and a display designer's knowledge of the task will necessarily inform and

constrain the design of a sonification.<sup>2</sup> A discussion of the types of tasks that users might undertake with sonifications, therefore, closely parallels the taxonomies of auditory displays described above.

#### *2.4.2.1 Monitoring*

Monitoring requires the listener to attend to a sonification over a course of time and to detect events (represented by sounds) and identify the meaning of the event in the context of the system's operation. These events are generally discrete and occur as the result of the attainment of some threshold in the system. Sonifications for monitoring tasks communicate the crossing of a threshold to the user, and they often require further (sometimes immediate) action in order for the system to operate properly (see the treatment of alerts and notifications above).

Kramer (1994) has described monitoring tasks as "template matching" in that the listener has a priori knowledge and expectations of a particular sound and its meaning. The acoustic pattern is already known, and the listener's task is to detect and identify the sound from a catalogue of known sounds. Consider a worker in an office environment that is saturated with intentional sounds from common devices, including telephones, fax machines, and computer interface sounds (e.g., email or instant messaging alerts). Part of the listener's task within such an environment is to monitor these devices. The alerting and notification sounds emitted from these devices facilitate that task in that they produce known acoustic patterns; the listener must hear and then match the pattern against the catalogue of known signals.

#### *2.4.2.2 Awareness of a process or situation*

Sonifications may sometimes be employed to promote the awareness of task-related processes or situations. Awareness-related task goals are different from monitoring tasks in that the sound coincides with or embellishes the occurrence of a process rather than simply indicating the crossing of a threshold that requires alerting. Whereas monitoring tasks may require action upon receipt of the message (e.g., answering a ringing phone or evacuating a building upon hearing a fire alarm), the sound signals that provide information regarding awareness may be less action-oriented and more akin to ongoing feedback regarding, or immersion in, task-related processes.

Nonspeech sounds like earcons and auditory icons, for example, have been used to enhance human-computer interfaces (see Brewster, 1997; Gaver, 1989). Typically, sounds are mapped to correspond to task-related processes in

---

<sup>2</sup> Human factors scientists have developed systematic methodologies for describing and understanding the tasks of humans in a man-machine system. Although an in-depth treatment of these issues is beyond the scope of this chapter, see Luczak (1997) for thorough coverage of task analysis purposes and methods.

the interface, such as scrolling, clicking, and dragging with the mouse, or deleting files, etc. Whereas the task that follows from monitoring an auditory display cannot occur in the absence of the sound signal (e.g., one can't answer a phone until it rings), the task-related processes in a computer interface can occur with or without the audio. The sounds are employed to promote awareness of the processes rather than to solely trigger some required response.

Similarly, soundscapes—ongoing ambient sonifications—have been employed to promote awareness of dynamic situations (a bottling plant, Gaver et al., 1991; financial data, Mauney & Walker, 2004; a crystal factory, Walker & Kramer, 2005). Although the soundscape may not require a particular response at any given time, it provides ongoing information about a situation to the listener.

#### *2.4.2.3 Data exploration*

Data exploration can entail any number of different subtasks ranging in purpose from holistic accounts of the entire data set to analytic tasks involving a single datum. Theoretical and applied accounts of visual graph and diagram comprehension have described a number of common tasks that are undertaken with quantitative data (see, for example, Cleveland & McGill, 1984; Friel, Curcio, & Bright, 2001; Meyer, 2000; Meyer, Shinar, & Leiser, 1997), and one can reasonably expect that the same basic categories of tasks will be required to explore data with auditory representations. The types of data exploration tasks described below are representative (but not necessarily comprehensive), and the chosen sonification approach may constrain the types of tasks that can be accomplished with the display and vice versa.

##### *2.4.2.3.1 Point estimation and point comparison*

Point estimation is an analytic listening task that involves extracting information regarding a single piece of information within a data set. Point estimation is fairly easily accomplished with data presented visually in a tabular format (Meyer, 2000), but data are quite likely to appear in a graphical format in scientific and popular publications (Zacks, Levy, Tversky, & Schiano, 2002). The extraction of information regarding a single datum, therefore, is a task that may need to be accomplished with an abstract (i.e., graphical) representation of the data rather than a table. Accordingly, researchers have begun to examine the extent to which point estimation is feasible with auditory representations of quantitative data such as auditory graphs. Smith and Walker (2005) performed a task analysis for point estimation with auditory graphs and determined that five steps were required to accomplish a point estimation task with sound. The listener must: 1) listen to the sonification; 2) determine in time when the datum of interest occurs; 3) upon identifying the datum of interest, estimate the magnitude of the quantity represented by the pitch of the tone; 4) compare this

magnitude to a baseline or reference tone (i.e., determine the scaling factor); and 5) report the value.

Point comparison, then, is simply comparing more than one datum; thus, point comparison involves performing point estimation twice (or more) and then using basic arithmetic operations to compare the two points. In theory, point comparison should be more difficult for listeners to perform accurately than point estimation, as listeners have twice as much opportunity to make errors, and there is the added memory component of the comparison task. Empirical investigations to date, however, have not examined point comparison tasks with sonifications.

#### 2.4.2.3.2 Trend identification

Trend identification is a more holistic listening task whereby a user attempts to identify the overall pattern of increases and decreases in quantitative data. Trend in a sonification closely parallels the notion of melodic contour in a piece of music. The listener may be concerned with global (overall) trend identification for data, or she/he may wish to determine local trends over a more narrow, specific time course within the sonification. Trend identification has been posited as a task for which the auditory system is particularly well-suited, and sound may be a medium wherein otherwise unnoticed patterns in data emerge for the listener.

#### 2.4.2.3.3 Identification of data structure

While the aforementioned tasks are primarily applicable to event-based sonification approaches, the goals of a model-based sonification user may be quite different. With model-based sonifications, the listener's task may involve identification of the overall structure of the model and thus the data. Through interactions with the virtual object, the listener hopes to extract information about the relationships within, and structure of, the data represented.

#### 2.4.2.3.4 Exploratory inspection

Occasionally, a user's task may be entirely exploratory in that she/he may wish to inspect or examine data with no a priori questions in mind. Kramer (1994) described exploratory tasks with sound as a less tractable endeavour than monitoring, because data exploration by its nature does not allow for an a priori, known catalogue of indicators. Still, the excellent temporal resolution of the auditory system and its pattern detection acuity make it a viable mode of data exploration, and the inspection of data with sound may reveal patterns and anomalies that were not perceptible in visual representations of the data.

#### *2.4.2.4 Dual task performance and multimodal tasking scenarios*

In many applications of sonifications, it is reasonable to assume that the human listener will likely have other auditory and/or visual tasks to perform in addition to working with the sonification. Surprisingly few studies to date have considered how the addition of a secondary task affects performance with sonifications. The few available studies are encouraging. Janata and Childs (2004) showed that sonifications aided a monitoring task with stock data, and the helpfulness of sound was even more pronounced when a secondary number-matching task was added. Peres and Lane (2005) found that while the addition of a visual monitoring task to an auditory monitoring task initially harmed performance of the auditory task, performance soon (i.e., after around 25 dual task trials) returned to pre-dual task levels. Brewster (1997) showed that the addition of sound to basic, traditionally visual interface operations enhanced performance of the tasks. Despite these encouraging results, a wealth of questions abound regarding the ability of listeners to use sonifications during concurrent visual and auditory tasks. Research to date has shed little light on the degree to which nonspeech audio interferes with concurrent processing of other sounds, including speech. The successful deployment of sonifications in real world settings will require a more solid base of knowledge regarding these issues.

## **2.5 Representation and Mappings**

Once the nature of the data and the task are determined, building a sonification involves mapping the data source(s) onto representational acoustic variables. This is especially true for parameter mapping techniques, but also applies, in a more general sense, to all sonifications. The mappings chosen by the display designer are an attempt to communicate information in each of the acoustic dimensions in use. It is important to consider how much of the intended “message” is received by the listener, and how close the perceived information matches the intended message.

### ***2.5.1 Semiotics: How acoustic perception takes on conceptual representation***

In some instances, sonification uses sound to portray an abstract representation of non-acoustic information; other circumstances might allow for an ecologically meaningful and valid sound to be employed in the display. In order to describe how to use sound as a sign or an indicator, however, it may be useful to first give a brief treatment of general approaches to how people come to assign meaning to sounds.

Semiotics is “the science of signs (and signals)” (Cuddon, 1991 p. 853). Clearly sonification aims to use sound to signify data or other information, and Pirhonen, Murphy, McAllister, and Yu (2006) have encouraged a semiotic

perspective in sound design. Empirical approaches, they argued, have been largely dominated by atheoretical, arbitrary sound design choices. Indeed the design space for sonifications is such that no study or series of studies could possibly make empirical comparisons of all combinations of sound manipulations. Pirhonen et al. argued for a semiotic approach to sound design that requires detailed use scenarios (describing a user and task) be presented to a design panel of experts or representative users. Such an approach seeks input regarding the most appropriate way to use sound as sign for particular users in a particular setting or context.

Kramer (1994) has described a representation continuum for sounds that ranges from analogic to symbolic (see Figure 2.2). At the extreme analogic end of the spectrum, the sound has the most direct and intrinsic relationship to its referent. Researchers have, for example, attempted to determine the extent to which the geometric shape of an object can be discerned by listening to the vibrations of physical objects that have been struck by mallets (Lakatos, McAdams, & Causse, 1997). At the symbolic end of the continuum, the referent may have an arbitrary or even random association with the sound employed by the display.

Keller and Stevens (2004) described the signal-referent relationships of environmental sounds with three categories: direct, indirect ecological, and indirect metaphorical. Direct relationships are those in which the sound is ecologically attributable to the referent. Indirect ecological relationships are those in which a sound that is ecologically associated with, but not directly attributable to, the referent is employed. Finally, indirect metaphorical relationships are those in which the sound signal is related to its referent only in some emblematic way.

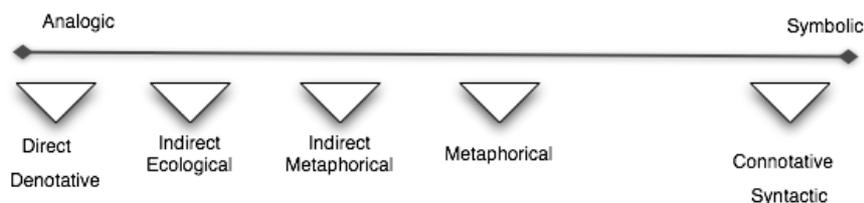


Figure 2.2: The analogic-symbolic representation continuum

### 2.5.2 *Semantic/iconic approach*

*Auditory icons*, mentioned earlier, are brief communicative sounds in an interface that bear an analogic relationship with the process they represent. In other words, the sound bears some ecological resemblance to the action or process (see Gaver, 1994; Kramer, 1994). This approach has also been called *sonic mapping* (Coward & Stevens, 2004). Auditory icons are appealing in that

the association between the sound and its intended meaning is more direct and should require little or no learning, but many of the actions and processes in a human-computer interface have no inherent auditory representation. For example, what should accompany a “save” action in a word processor? How can that sound be made distinct from a similar command, such as “save as”? *Earcons*, on the other hand, use sounds as symbolic representations of actions or processes; the sounds have no ecological relationship to their referent (see Blattner, Sumikawa, & Greenberg, 1989; Kramer, 1994). Earcons are made by systematically manipulating the pitch, timbre, and rhythmic properties of sounds to create a structured set of non-speech sounds that can be used to represent any object or concept through an arbitrary mapping of sound to meaning. Repetitive or related sequences or motifs may be employed to create “families” of sounds that map to related actions or processes. While earcons can represent virtually anything, making them more flexible than auditory icons, a trade-off exists in that the abstract nature of earcons may require longer learning time or even formal training in their use. Walker (2006) has discussed a new type of interface sound, the *spearcon*, which is intended to overcome the shortcomings of both auditory icons and earcons. Spearcons are created by speeding up a spoken phrase to the point where it is no longer recognizable as speech, and as such can represent anything (like earcons can), but are non-arbitrarily mapped to their concept (like auditory icons). The main point here is that there are tradeoffs when choosing how to represent a concept with a sound, and the designer needs to make explicit choices with the tradeoffs in mind.

### **2.5.3 Choice of display dimension**

When creating a more typical parameter-mapped sonification, such as representing rainfall and average daily temperature over the past year, the issues of mapping, polarity, and scaling are crucial (Walker, 2002, in press; Walker & Kramer, 2004).

#### **2.5.3.1 Data-to-display Mapping**

In sonification it matters which specific sound dimension is chosen to represent a given data dimension. This is partly because there seems to be some agreement among listeners about what sound attributes are good (or poor) at representing particular data dimensions. For example, pitch is generally good for representing temperature, whereas tempo is not as effective (Walker, 2002). It is also partly because some sound dimensions (e.g., loudness) are simply not very effective in auditory displays for practical design reasons (Neuhoff, Kramer, & Wayand, 2002). Walker has recently evaluated mappings between ten conceptual data dimensions (e.g., temperature, pressure, danger) and three perceptual/acoustic dimensions (pitch, tempo, and spectral brightness), in an effort to determine which sounds should be used to represent a given type of data (see also Walker, 2002, in press). This type of research will need to be extended

to provide designers with guidance about mapping choices. In turn, sonification designers need to be aware that not all mappings are created equal, and must use a combination of empirically-derived guidelines and usability testing to ensure the message they are intending to communicate is being received by the listener. In addition to those already discussed, guidelines for mappings from a variety of sources should be consulted (e.g., Bonebright, Nees, Connerley, & McCain, 2001; L. M. Brown, Brewster, Ramloll, Burton, & Riedel, 2003; Flowers & Hauer, 1995; Neuhoff & Heller, 2005; Smith & Walker, 2005; Walker, 2002).

#### *2.5.3.2 Mapping Polarity*

Sonification success also requires an appropriate polarity for the data-to-display mappings. For example, listeners might agree that pitch should increase in order to represent increasing temperature (a positive mapping polarity, Walker, 2002), while at the same time feel that pitch should decrease in order to represent increasing size (a *negative* polarity). The issue of polarity is not typically an issue for visual displays, but it can be very important in auditory representations ranging from helicopter warning sounds (Edworthy et al., 2004) to interfaces for the visually impaired (Walker & Lane, 2001). Walker (2002, in press) lists the preferred polarities for many mappings, and points out that performance is actually impacted with polarities that do not match listener expectancies. Again, a mixture of guidelines and testing are important to ensure that a sonification is in line with what listeners anticipate.

#### *2.5.3.3 Scaling*

Once an effective mapping and polarity has been chosen, it is important to determine how much change in, say, the pitch of a sound is used to convey a given change in, for example, temperature. Matching the data-to-display scaling function to the listener's internal conceptual scaling function between pitch and temperature is critical if the sonification is to be used to make accurate comparisons and absolute or exact judgments of data values, as opposed to simple trend estimations. This is a key distinction between sonifications and warnings or trend monitoring sounds. Again, Walker (2002, in press) has empirically determined scaling factors for several mappings, in both positive and negative polarities. Such values begin to provide guidance about how different data sets would be represented most effectively. However, it is important not to over-interpret the exact exponent values reported in any single study, to the point where they are considered "*the*" correct values for use in all cases. As with any performance data that are used to drive interface guidelines, care must always be taken to avoid treating the numbers as components of a design recipe. Rather, they should be treated as guidance, at least until repeated measurements and continued application experience converge toward a clear value or range.

Beyond the somewhat specific scaling factors discussed to this point, there are some practical considerations that relate to scaling issues. Consider, for

example, using frequency changes to represent average daily temperature data that ranges from 0-30° Celsius. The temperature data *could* be scaled to fill the entire hearing range (best case, about 20 Hz to 20,000 Hz); but a much more successful approach might be to scale the data to the range where hearing is most sensitive, say between 1000-5000 Hz. Another approach would be to base the scaling on a musical model, where the perceptually equal steps of the notes on a piano provide a convenient scale. For this reason, computer music approaches to sonification, including mapping data onto MIDI notes, have often been employed. Limiting the range of notes has often been recommended (e.g., using only MIDI notes 35-100, L. M. Brown et al., 2003). Even in that case, the designer has only 65 display points to use to represent whatever data they may have. Thus, the granularity of the scale is limited. For the daily temperature data that may be sufficient, but other data sets may require more precision. A designer may be forced to “round” the data values to fit the scale, or alternatively employ “pitch bending” to play a note at the exact pitch required by the data. This tends to take away from the intended musicality of the approach. Again, this is a tradeoff that the designer needs to consider. Some software (e.g., the Sonification Sandbox, Walker & Cothran, 2003; Walker & Lowey, 2004) provides both rounding and exact scaling options, so the one that is most appropriate can be used, given the data and the tasks of the listener.

#### *2.5.3.4 Concurrent presentation of multiple data streams/series*

Many data analysis tasks require the comparison of values from more than one data source presented concurrently. This could be daily temperatures from different cities, or stock prices from different stocks. The general theory invoked in this situation is auditory streaming (Bregman, 1990). In some cases (for some tasks), it is important to be able to perceptually separate or segregate the different city data, whereas in other cases it is preferable for the two streams of data to fuse into a perceptual whole. Bregman (1990) discusses what acoustic properties support or inhibit stream segregation. Briefly, differences in timbre (often achieved by changing the musical instrument; see Cusack & Roberts, 2000, for a report on the utility of timbre in stream segregation) and spatial location (or stereo panning) are parameters that sonification designers can often use simply and effectively (see also Bonebright et al., 2001; L. M. Brown et al., 2003). McGookin and Brewster (2004) have shown that, while increasing the number of concurrently presented earcons decreases their identifiability, such problems can be somewhat overcome by introducing timbre and onset differences. Pitch is another attribute that can be used to segregate streams, but in sonification pitch is often dynamic (being used to represent changing data values), so it is a less controllable and less reliable attribute for manipulating segregation.

#### 2.5.3.5 Context

*Context* refers to the purposeful addition of non-signal information to a display (Smith & Walker, 2005; Walker & Nees, 2005a). In visual displays, additional information such as axes and tick marks can increase readability and aid perception by enabling more effective top-down processing (Bertin, 1983; Tufte, 1990). A visual graph without context cues (e.g., no axes or tick marks) provides no way to estimate the value at any point. The contour of the line provides some *incidental* context, which might allow an observer to perform a trend analysis (rising versus falling), but the accurate extraction of a specific value (i.e., a point estimation task) is impossible.

Even sonifications that make optimal use of mappings, polarities, and scaling factors need to include contextual cues equivalent to axes, tick marks and labels, so the listener can perform the interpretation tasks. Recent work (Nees & Walker, 2006; Smith & Walker, 2005) has shown that even for simple sonifications, the addition of some kinds of context cues can provide useful information to users of the display. For example, simply adding a series of clicks to the display can help the listener keep track of the time better, which keeps their interpretation of the graph values more “in phase” (see also Bonebright et al., 2001; Flowers et al., 1997; Gardner, Lundquist, & Sahyun, 1996). Smith and Walker (2005) showed that when the clicks played at a rate that was twice the rate of the sounds representing the data, the two sources of information combined like the major and minor tick marks on the x-axis of a visual graph. The addition of a repeating reference tone that signified the maximum value of the data set provided dramatic improvements in the attempts by listeners to estimate exact data values, whereas a reference tone that signified the starting value of the data did not improve performance. Thus, it is clear that adding context cues to auditory graphs can play the role that x- and y-axes play in visual graphs, but not all implementations are equally successful. Researchers have only scratched the surface of possible context cues and their configurations, and we need to implement and validate other, perhaps more effective, methods (see, e.g., Nees & Walker, 2006).

## 2.6 Limiting Factors for Sonification: Aesthetics, Individual Differences, and Training

Although future research should shed light on the extent to which particular tasks and data sets are amenable to representation with sound, the major limiting factors in the deployment of sonifications have been, and will continue to be, the perceptual and information processing capabilities of the human listener.

### **2.6.1 *Aesthetics and musicality***

Edworthy (1998) aptly pointed out the independence of display performance and aesthetics. While sound may aesthetically enhance a listener's interaction with a system, performance may not necessarily be impacted by the presence or absence of sound. Questions of aesthetics and musicality remain open in the field of sonification. The use of musical sounds (as opposed to pure sine wave tones, etc.) has been recommended because of the ease with which musical sounds are perceived (L. M. Brown et al., 2003), but it remains to be seen whether the use of musical sounds such as those available in MIDI instrument banks affords performance improvements over less musical, and presumably less aesthetically desirable, sounds. Although the resolution of issues regarding aesthetics and musicality is clearly relevant, it nevertheless remains advisable to design aesthetically pleasing (i.e., musical, etc.) sonifications to the extent possible while still conveying the intended message. Kramer (1994) identified listener annoyance as a potential factor that would deter the use of auditory displays; thus a designer should aim to avoid annoyance as much as possible.

### **2.6.2 *Individual differences and training***

The capabilities, limitations, and experiences of listeners, as well as transient states (like mood and level of fatigue) will all impact performance outcomes with auditory displays. Surprisingly little is known about the impact of between- and within-individual differences on auditory display outcomes. Understanding individual differences in perceptual, cognitive, and musical abilities of listeners will inform the design of sonifications in several important ways. First, by understanding ranges in individual difference variables, a designer can, where required, build a display that accommodates all users in a given context (e.g., universal design, see Iwarsson & Stahl, 2003). Furthermore, in situations where only optimal display users are desirable, understanding the relevance and impact of individual difference variables will allow for the selection of display operators whose capabilities will maximize the likelihood of success with the display. Finally, the extent to which differences in training and experience with sonifications affects performance with the displays is a topic deserving further investigation.

#### **2.6.2.1 *Perceptual capabilities of the listener***

A treatment of theoretical issues relevant to sonification would be remiss not to mention those characteristics of the human listener that impact comprehension of auditory displays. The fields of psychoacoustics and basic auditory perception (see Chapters 3 and 4 in this volume) have offered critical insights for the design and application of sonifications. As Walker and Kramer (2004) pointed out, these fields have contributed a widely accepted vocabulary

and methodology to the study of sound perception, as well as a foundation of knowledge that is indispensable to the study of sonification.

Detection is of course a crucial first consideration for auditory display design. The listener must be able to hear the sound(s) in the environment in which the display is deployed. Psychoacoustic research has offered insights into minimum thresholds (e.g., see Hartmann, 1997; Licklider, 1951), and masking theories offer useful predictions regarding the detectability of a given acoustic signal in noise (for a discussion, see Mulligan, McBride, & Goodman, 1984; C. S. Watson & Kidd, 1994). Empirical data for threshold and masking studies, however, are usually gathered in carefully controlled settings with minimal stimulus uncertainty. As Watson and Kidd (1994) and others (e.g., Mulligan et al., 1984; Walker & Kramer, 2004) point out, such data may provide apt descriptions of auditory capabilities but poor guidelines for auditory display design. The characteristics of the environment in which a display operates may differ drastically from the ideal testing conditions and pure tone stimuli of psychophysical experiments. As a result, Watson and Kidd suggested that ecologically valid testing conditions for auditory displays should be employed to establish real-world guidelines for auditory capabilities. Furthermore, recent work has drawn attention to the phenomenon of informational masking, whereby sounds that theoretically should *not* be masked in the peripheral hearing mechanism (i.e., the cochlea) are indeed masked, presumably at higher levels in the auditory system (see Durlach et al., 2003). Clearly, the seemingly straightforward requirement of detectability for auditory displays warrants a careful consideration of the display's user as well as the environments and apparatus (headphones, speakers, etc.) with which the display will be implemented.

Beyond basic knowledge of the detectability of sound, auditory display designers should be aware of the psychophysical limitations on judgements of discrimination (e.g., just-noticeable differences, etc.) and identification of sounds. Again, however, the extant data regarding average absolute discrimination or identification performance in extremely controlled conditions may be misleading if one expects the same level of performance with different stimuli (i.e., sounds that are not pure tones) in less controlled, non-laboratory environments. Sonification researchers can and should, however, actively borrow from and adapt the knowledge and methods of psychoacousticians. For example, Sandor and Lane (2003) introduced the term *mappable difference* to describe the absolute error in response accuracy (i.e., criterion) one must allow for in order to achieve a given proportion of accurate responses for a point estimation sonification task. Such a metric also allowed them to identify the number of distinct values that could be represented with a given proportion of accuracy for their chosen scales. Such innovative approaches that combine the methods and tools of psychoacoustics and perception with the real world stimuli and applications of auditory display designers may be the best approach to

understanding how to maximize information transmission with auditory displays by playing to the strengths of the human perceiver.

#### *2.6.2.2 Cognitive abilities of the listener*

Researchers have posited roles for a number of cognitive abilities in the comprehension of visual displays, including spatial abilities (Trickett & Trafton, 2006), domain or content knowledge and graph-reading skill (Shah, 2002), and working memory (Toth & Lewis, 2002). The role of such cognitive abilities in the comprehension of sonifications and auditory stimuli in general, however, remains relatively unexplored. The few studies that have examined relationships between cognitive abilities and auditory perception have found results that suggest cognitive individual differences will impact auditory display performance. Payne (2003), for example, found relationships between working memory and pitch discrimination. Walker and Mauney (2004) found that spatial reasoning ability predicts some variance in performance with auditory graphs. More research is needed to determine the full array of cognitive factors contributing to auditory display performance, and the extent to which such cognitive abilities can be accurately assessed and used to predict performance.

Additionally, questions regarding the cognitive representations formed and used by auditory display listeners remain virtually untouched. For example, if, as Kramer (1994) argued, sonification monitoring tasks employ template matching processes, then what are the properties of the stored templates and how are they formed? In the case of auditory graphs, do people attempt to translate the auditory stimulus into a more familiar visual mental representation? Anecdotal evidence reported by Flowers (1995) suggested that listeners were indeed inclined to draw visual representations of auditory graphs on scrap paper during testing. Though sonification research tends to shy away from basic and theoretical science in favor of more applied lines of research, studies leading to better accounts of the cognitive representations of sonifications would favourably inform display design.

#### *2.6.2.3 Musical abilities of the listener*

For many years, researchers predicted and anticipated that musicians would outperform non-musicians on tasks involving auditory displays. Musical experience and ability, then, have been suggested as individual level predictors of performance with auditory displays, but research has generally found weak to non-existent correlations between musical experience and performance with auditory displays. One plausible explanation for the lack of relationship between musicianship and auditory display performance is the crude nature of self-report metrics of musical experience, which are often the yardstick for describing the degree to which a person has musical training. A person could have had many years of musical experience as child, yet that person could be many years removed from their musical training and exhibit no more musical *ability* than

someone who received no formal training. A more fruitful approach to the measurement of musicianship in the future may be to develop brief, reliable, and valid measure of musical ability for diagnostic purposes in research (e.g., Edwards, Challis, Hankinson, & Pirie, 2000), along the lines of research in musical abilities by Seashore and others (e.g., A. W. Brown, 1928; Cary, 1923; Seashore, Lewis, & Saetveit, 1960).

Although the predictive value of individual differences in musical ability is worthy of further study and differences between musicians and nonmusicians have been reported (e.g., Neuhoff & Wayand, 2002; Sandor & Lane, 2003), the ultimate contribution of musical ability to performance with auditory displays may be minor. Watson and Kidd (1994) suggested that the auditory perceptual abilities of the worst musicians are likely better than the abilities of the worst non-musicians, but the best non-musicians likely have auditory perceptual abilities on par with the best musicians.

#### *2.6.2.4 Training*

Sonification offers a novel approach to information representation, and this novelty stands as a potential barrier to the success of the display unless the user can be thoroughly and efficiently acclimated to the meaning of the sounds being presented. Visual information displays owe much of their success to their pervasiveness as well as to users' formal education and informal experience at deciphering their meanings. Graphs, a basic form of visual display, are incredibly pervasive in print media (see Zacks et al., 2002), and virtually all children are taught how to read graphs from a very young age in formal education settings. Complex auditory displays currently are not pervasive, and users are not taught how to comprehend auditory displays as part of a standard education. This problem can be partially addressed by exploiting the natural analytic prowess and intuitive, natural meaning-making processes of the auditory system (see Gaver, 1993), but training will likely be necessary even when ecological approaches to sound design are pursued.

To date, little attention has been paid to the issue of training sonification users. Empirical findings suggesting that sonifications can be effective are particularly encouraging considering that the majority of these studies sampled naïve users who had presumably never listened to sonifications before entering the lab. For the most part, information regarding performance ceilings for sonifications remains speculative, as few or no studies have examined the role of extended training in performance.

As Watson and Kidd (1994) suggested, many populations of users may be unwilling to undergo more than nominally time-consuming training programs, but research suggests that even brief training for sonification users offers benefits. Smith and Walker (2005) showed that brief training for a point estimation task (i.e., naming the Y axis value for a given X axis value in an auditory graph) resulted in better performance than no training, while Walker

and Nees (2005b) further demonstrated that a brief training period (around 20 min) can reduce performance error by 50% on a point estimation sonification task. Recent and ongoing work is examining exactly what kinds of training methods are most effective for different classes of sonifications (e.g., Walker & Nees, 2005c).

## **2.7 Conclusions: Toward a Cohesive Theoretical Account of Sonification**

Despite a tremendous growth in the interest in and feasibility of sonification over the past two decades, the field as a whole continues to exist in a virtual theoretical vacuum whereby design decisions are often made arbitrarily or at the best judgement of an informed designer (Edworthy, 1998; Frauenberger & Stockman, 2006; Pirhonen et al., 2006). Although the field of auditory display has benefited tremendously from multidisciplinary approaches in research and practice, this same diversity has likely been an obstacle to the formation of a unified, cohesive theoretical account of sound as an information display modality. To date, few theories or models of human interaction with auditory displays exist. Over 15 years ago, Edwards (1989) adapted the Keystroke Level Model (see Card & Moran, 1980) of user interaction with visual interfaces to describe the behavior of auditory interface users. Edwards explained, “This work represents a first step towards expanding models of human-computer interactions to include auditory interactions” (p. 575). Unfortunately, few researchers followed his lead, but it seems inevitable that those in the field of sonification will need to develop explanatory models in order to achieve a cohesive theoretical stance. As Edwards pointed out, the development of new models or the expansion of existing models of human interaction with information systems to include auditory displays will benefit twofold: 1) In research, models of human interaction with auditory displays will provide testable hypotheses that will guide a systematic, programmatic approach to auditory display research, and 2) In practice, auditory display designers will be able to turn to models for basic guidelines.

Furthermore, a definitive theoretical description of sonification will require the assimilation of knowledge from the past several decades across a number of fields. Sonification should not “reinvent the wheel” in instances where past research from other fields (acoustics, auditory perception, etc.), has already offered clear insights. For example, Williams and Aiken (1975; Williams & Aiken, 1977) conducted pattern classification studies 30 years ago that offered an early attempt to systematically map auditory dimensions to visual representations of quantitative data using changes in frequency over time. Although they did not call their stimuli auditory graphs, they were remarkably graphical in nature, yet these studies seem to go unnoticed by auditory graph researchers. Over 50 years ago, Pollack (1952; Pollack & Ficks, 1954)

innovatively applied information theory principles to evaluate auditory displays as communicative tools, but all these years later many current investigations of sonification seem to proceed uninformed by *any* guiding theory.

[some final “ease out the door” remarks...]

#### References

- Barrass, S. (1997). Auditory information design. Unpublished Dissertation, Australian National University.
- Barrass, S. (2005). A perceptual framework for the auditory display of scientific data. *ACM Transactions on Applied Perception*, 2(4), 389-492.
- Bertin, J. (1983). *Semiology of graphics* (W. J. Berg, Trans.). Madison, Wisconsin: The University of Wisconsin Press.
- 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.
- Bonebright, T. L., Nees, M. A., Connerley, T. T., & McCain, G. R. (2001). Testing the effectiveness of sonified graphs for education: A programmatic research project. *Proceedings of the International Conference on Auditory Display (ICAD2001)* (pp. 62-66), Espoo, Finland.
- Bovermann, T., Hermann, T., & Ritter, H. (2006). Tangible data scanning sonification model. *Proceedings of the 12th International Conference on Auditory Display (ICAD06)* (pp. 77-82), London, UK.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Brewster, S. (1997). Using non-speech sound to overcome information overload. *Displays*, 17, 179-189.
- Brewster, S., & Murray, R. (2000). Presenting dynamic information on mobile computers. *Personal Technologies*, 4(4), 209-212.
- Brown, A. W. (1928). The reliability and validity of the Seashore Tests of Musical Talent. *Journal of Applied Psychology*, 12, 468-476.
- Brown, L. M., Brewster, S., & Riedel, B. (2002). Browsing modes for exploring sonified line graphs. *Proceedings of the 16th British HCI Conference*, London, UK.
- Brown, L. M., & Brewster, S. A. (2003). Drawing by ear: Interpreting sonified line graphs. *Proceedings of the International Conference on Auditory Display (ICAD2003)* (pp. 152-156), Boston, MA.
- Brown, L. M., Brewster, S. A., Ramloll, R., Burton, M., & Riedel, B. (2003). Design guidelines for audio presentation of graphs and tables.

- Proceedings of the International Conference on Auditory Display (ICAD2003)* (pp. 284-287), Boston, MA.
- Brown, M. L., Newsome, S. L., & Glinert, E. P. (1989). An experiment into the use of auditory cues to reduce visual workload. *Proceedings of the ACM CHI 89 Human Factors in Computing Systems Conference (CHI 89)* (pp. 339-346).
- Buxton, W. (1989). Introduction to this special issue on nonspeech audio. *Human-computer Interaction*, 4, 1-9.
- Card, S. K., & Moran, T. P. (1980). The keystroke-level model for user performance time with interactive systems. *Communications of the ACM*, 23(7), 396-410.
- Cary, H. (1923). Are you a musician? Professor Seashore's specific psychological tests for specific musical abilities. *Scientific American*, 326-327.
- Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, 79(387), 531-554.
- Coward, S. W., & Stevens, C. J. (2004). Extracting meaning from sound: Nomic mappings, everyday listening, and perceiving object size from frequency. *The Psychological Record*, 54, 349-364.
- Cuddon, J. A. (1991). *Dictionary of Literary Terms and Literary Theory* (3rd ed.). New York: Penguin Books.
- Cusack, R., & Roberts, B. (2000). Effects of differences in timbre on sequential grouping. *Perception & Psychophysics*, 62(5), 1112-11120.
- de Campo, A. (2006). Data sonification design space map. Unpublished manuscript.
- 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.
- Durlach, N. I., Mason, C. R., Kidd, G., Arbogast, T. L., Colburn, H. S., & Shinn-Cunningham, B. (2003). Note on informational masking. *Journal of the Acoustical Society of America*, 113(6), 2984-2987.
- Edwards, A. D. N. (1989). Modelling blind users' interactions with an auditory computer interface. *International Journal of Man-Machine Studies*, 30(575-589).
- Edwards, A. D. N., Challis, B. P., Hankinson, J. C. K., & Pirie, F. L. (2000). Development of a standard test of musical ability for participants in auditory interface testing. *Proceedings of the International Conference on Auditory Display (ICAD 2000)*, Atlanta, GA.
- Edworthy, J. (1998). Does sound help us to work better with machines? A commentary on Rautenberg's paper 'About the importance of auditory

- alarms during the operation of a plant simulator'. *Interacting with Computers*, 10, 401-409.
- Edworthy, J., & Hellier, E. (2006). Complex nonverbal auditory signals and speech warnings. In M. S. Wogalter (Ed.), *Handbook of Warnings* (pp. 199-220). Mahwah, NJ: Lawrence Erlbaum.
- Edworthy, J., Hellier, E. J., Aldrich, K., & Loxley, S. (2004). Designing trend-monitoring sounds for helicopters: Methodological issues and an application. *Journal of Experimental Psychology: Applied*, 10(4), 203-218.
- Fitch, W. T., & Kramer, G. (1994). Sonifying the body electric: Superiority of an auditory over a visual display in a complex, multivariate system. In G. Kramer (Ed.), *Auditory display: Sonification, audification, and auditory interfaces* (pp. 307-326). Reading, MA: Addison-Wesley.
- 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 assess characteristics of numeric data: Are we too visuocentric? *Behavior Research Methods, Instruments & 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 & 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.
- Franklin, K. M., & Roberts, J. C. (2004). A path based model for sonification. *Proceedings of the Eighth International Conference on Information Visualization (IV '04)* (pp. 865-870).
- Frauenberger, C., & Stockman, T. (2006). Patterns in auditory menu design. *Proceedings of the 12th International Conference on Auditory Display (ICAD06)* (pp. 141-147), London, UK.
- Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional applications [Electronic version]. *Journal for Research in Mathematics*, 32(2), 124-159.
- Frynsinger, S. P. (2005). A brief history of auditory data representation to the 1980s. *Proceedings of the International Conference on Auditory Display (ICAD 2005)*, Limerick, Ireland.
- Gardner, J. A., Lundquist, R., & Sahyun, S. (1996). TRIANGLE: A practical application of non-speech audio for imparting information. *Proceedings of the International Conference on Auditory Display* (pp. 59-60), San Francisco, CA.

- Garner, W. R., & Gottwald, R. L. (1968). The perception and learning of temporal patterns. *The Quarterly Journal of Experimental Psychology*, 20(2).
- Gaver, W. W. (1989). The SonicFinder: An interface that uses auditory icons. *Human-Computer Interaction*, 4(1), 67-94.
- Gaver, W. W. (1993). What in the world do we hear? An ecological approach to auditory event perception. *Ecological Psychology*, 5(1), 1-29.
- 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 ACM Conference on Human Factors in Computing Systems CHI'91*, New Orleans.
- Global music - The world by ear. (2006). [Concert]. London, UK: International Conference on Auditory Display (ICAD06).
- Haas, E., & Edworthy, J. (2006). An introduction to auditory warnings and alarms. In M. S. Wogalter (Ed.), *Handbook of Warnings* (pp. 189-198). Mahwah, NJ: Lawrence Erlbaum
- Hermann, T., & Hunt, A. (2005). An introduction to interactive sonification. *IEEE Multimedia*, 12(2), 20-24.
- Iwarsson, S., & Stahl, A. (2003). Accessibility, usability, and universal design--positioning and definition of concepts describing person-environment relationships. *Disability and Rehabilitation*, 25(2), 57-66.
- Janata, P., & Childs, E. (2004). Marketbuzz: Sonification of real-time financial data. *Proceedings of the Tenth Meeting of the International Conference on Auditory Display (ICAD04)*, Sydney, Australia.
- Johannsen, G. (2004). Auditory displays in human-machine interfaces. *Proceedings of the IEEE*, 92(4), 742-758.
- Keller, P., & Stevens, C. (2004). Meaning from environmental sounds: Types of signal-referent relations and their effect on recognizing auditory icons. *Journal of Experimental Psychology: Applied*, 10(1), 3-12.
- Kortum, P., Peres, S. C., Knott, B., & Bushey, R. (2005). The effect of auditory progress bars on consumer's estimation of telephone wait time. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting* (pp. 628-632), Orlando, FL.
- Kramer, G. (1994). 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., 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).

- Lakatos, S., McAdams, S., & Causse, R. (1997). The representation of auditory source characteristics: Simple geometric form. *Perception & Psychophysics*, 59(8), 1180-1190.
- Levitin, D. J. (1999). Memory for musical attributes. In P. Cook (Ed.), *Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics*. (pp. 209-227). Cambridge, MA: MIT Press.
- Listening to the mind listening: Concert of sonifications at the Sydney Opera House. (2004). [Concert]. Sydney, Australia: International Conference on Auditory Display (ICAD04).
- Luczak, H. (1997). Task analysis. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (2nd ed., pp. 340-416). New York: Wiley.
- Mauney, B. S., & Walker, B. N. (2004). Creating functional and livable soundscapes for peripheral monitoring of dynamic data. *Proceedings of the 10th International Conference on Auditory Display (ICAD04)*, Sydney, Australia.
- McAdams, S., & Bigand, E. (1993). *Thinking in sound: the cognitive psychology of human audition*. Oxford: Oxford University Press.
- McGookin, D. K., & Brewster, S. A. (2004). Understanding concurrent earcons: Applying auditory scene analysis principles to concurrent earcon recognition. *Acm Transactions on Applied Perceptions*, 1, 130-150.
- Meyer, J. (2000). Performance with tables and graphs: Effects of training and a visual search model. *Ergonomics*, 43(11), 1840-1865.
- Meyer, J., Shinar, D., & Leiser, D. (1997). Multiple factors that determine performance with tables and graphs. *Human Factors*, 39(2), 268-286.
- Moore, B. C. J. (1997). *An introduction to the psychology of hearing* (4th ed.). San Diego, Calif.: Academic Press.
- Mulligan, B. E., McBride, D. K., & Goodman, L. S. (1984). A design guide for nonspeech auditory displays: Naval Aerospace Medical Research Laboratory Technical Report.
- Nees, M. A., & Walker, B. N. (2006). Relative intensity of auditory context for auditory graph design. *Proceedings of the Twelfth International Conference on Auditory Display (ICAD06)* (pp. 95-98), London, UK.
- Neuhoff, J. G., & Heller, L. M. (2005). One small step: Sound sources and events as the basis for auditory graphs. *Proceedings of the Eleventh Meeting of the International Conference on Auditory Display*, Limerick, Ireland.
- Neuhoff, J. G., Kramer, G., & Wayand, J. (2002). Pitch and loudness interact in auditory displays: Can the data get lost in the map? *Journal of Experimental Psychology: Applied*, 8(1), 17-25.
- Neuhoff, J. G., & Wayand, J. (2002). Pitch change, sonification, and musical expertise: Which way is up? *Proceedings of the International Conference on Auditory Display* (pp. 351-356), Kyoto, Japan.

- Payne, T. W. (2003). Working memory capacity and pitch discrimination: Unpublished doctoral dissertation, Georgia Institute of Technology, Atlanta, GA.
- Peres, S. C., & Lane, D. M. (2005). Auditory graphs: The effects of redundant dimensions and divided attention. *Proceedings of the International Conference on Auditory Display (ICAD 2005)* (pp. 169-174), Limerick, Ireland.
- Pirhonen, A., Murphy, E., McAllister, g., & Yu, W. (2006). Non-speech sounds as elements of a use scenario: A semiotic perspective. *Proceedings of the 12th International Conference on Auditory Display (ICAD06)*, London, UK.
- Pollack, I. (1952). The information of elementary auditory displays. *Journal of the Acoustical Society of America*, 24(6), 745-749.
- Pollack, I., & Ficks, L. (1954). Information of elementary multidimensional auditory displays. *Journal of the Acoustical Society of America*, 26, 155-158.
- Quinn, M. (2001). Research set to music: The climate symphony and other sonifications of ice core, radar, DNA, seismic, and solar wind data. *Proceedings of the 7th International Conference on Auditory Display (ICAD01)*, Espoo, Finland.
- Quinn, M. (2003). For those who died: A 9/11 tribute. *Proceedings of the 9th International Conference on Auditory Display*, Boston, MA.
- 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.
- Sanderson, P. M. (in press). Designing for attention with sound: Challenges and extensions to Ecological Interface Design. *Human Factors*, 48(4).
- Sandor, A., & Lane, D. M. (2003). Sonification of absolute values with single and multiple dimensions. *Proceedings of the 2003 International Conference on Auditory Display (ICAD03)* (pp. 243-246), Boston, MA.
- Seashore, C. E., Lewis, D., & Saetveit, J. G. (1960). *Seashore Measures of Musical Talents* (Revised 1960 ed.). New York: The Psychological Corp.
- Shah, P. (2002). Graph comprehension: The role of format, content, and individual differences. In M. Anderson, B. Meyer & P. Olivier (Eds.), *Diagrammatic Representation and Reasoning* (pp. 173-185). New York: Springer.
- Shannon, C. E. (1998/1949). Communication in the presence of noise. *Proceedings of the IEEE*, 86(2), 447-457.
- Smith, D. R., & Walker, B. N. (2005). Effects of auditory context cues and training on performance of a point estimation sonification task. *Journal of Applied Cognitive Psychology*, 19(8), 1065-1087.

- 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.
- Spence, C., & Driver, J. (1997). Audiovisual links in attention: Implications for interface design. In D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics Vol. 2: Job Design and Product Design* (pp. 185-192). Hampshire: Ashgate Publishing.
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 13(2684), 677-680.
- Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects*. New York: Wiley.
- Stokes, A., Wickens, C. D., & Kite, K. (1990). *Display Technology: Human Factors Concepts*. Warrendale, PA: Society of Automotive Engineers.
- Toth, J. A., & Lewis, C. M. (2002). The role of representation and working memory in diagrammatic reasoning and decision making. In M. Anderson, B. Meyer & P. Olivier (Eds.), *Diagrammatic Representation and Reasoning* (pp. 207-221). New York: Springer.
- Trickett, S. B., & Trafton, J. G. (2006). Toward a comprehensive model of graph comprehension: Making the case for spatial cognition. *Proceedings of the Fourth International Conference on the Theory and Application of Diagrams (DIAGRAMS 2006)*, Stanford University, USA.
- Tufte, E. R. (1990). *Envisioning information*. Cheshire, Connecticut: Graphics Press.
- Walker, B. N. (2002). Magnitude estimation of conceptual data dimensions for use in sonification. *Journal of Experimental Psychology: Applied*, 8, 211-221.
- Walker, B. N. (in press). Consistency of magnitude estimations with conceptual data dimensions used for sonification. *Journal of Applied Cognitive Psychology*.
- Walker, B. N., & Cothran, J. T. (2003). Sonification Sandbox: A graphical toolkit for auditory graphs. *Proceedings of the International Conference on Auditory Display (ICAD2003)* (pp. 161-163), Boston, MA.
- Walker, B. N., & Kramer, G. (1996). 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. (2004). Ecological psychoacoustics and auditory displays: Hearing, grouping, and meaning making. In J. Neuhoff (Ed.), *Ecological psychoacoustics* (pp. 150-175). New York: Academic Press.
- Walker, B. N., & Kramer, G. (2005). Mappings and metaphors in auditory displays: An experimental assessment. *ACM Transactions on Applied Perception*, 2(4), 407-412.

- Walker, B. N., & Lane, D. M. (2001). Psychophysical scaling of sonification mappings: A comparison of visually impaired and sighted listeners. *Proceedings of the 7th International Conference on Auditory Display* (pp. 90-94), Espoo, Finland.
- Walker, B. N., & Lowey, M. (2004). Sonification Sandbox: A graphical toolkit for auditory graphs. *Proceedings of the Rehabilitation Engineering & Assistive Technology Society of America (RESNA) 27th International Conference*, Orlando, FL.
- Walker, B. N., & Mauney, L. M. (2004). Individual differences, cognitive abilities, and the interpretation of auditory graphs. *Proceedings of the International Conference on Auditory Display (ICAD2004)*, Sydney, Australia.
- Walker, B. N., Nance, A., & Lindsay, J. (2006). Spearcons: Speech-based earcons improve navigation performance in auditory menus. *Proceedings of the International Conference on Auditory Display (ICAD06)* (pp. 63-68), London, UK.
- Walker, B. N., & Nees, M. A. (2005a). An agenda for research and development of multimodal graphs. *Proceedings of the International Conference on Auditory Display (ICAD2005)* (pp. 428-432), Limerick, Ireland.
- Walker, B. N., & Nees, M. A. (2005b). Brief training for performance of a point estimation task sonification task. *Proceedings of the International Conference on Auditory Display (ICAD2005)* (pp. TBD.), Limerick, Ireland.
- Walker, B. N., & Nees, M. A. (2005c). Conceptual versus perceptual training for auditory graphs. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting* (pp. 1598-1601), Orlando, FL.
- Watson, C. S., & Kidd, G. R. (1994). Factors in the design of effective auditory displays. *Proceedings of the International Conference on Auditory Display (ICAD1994)*, Sante Fe, NM.
- Watson, M. (2006). Scalable earcons: Bridging the gap between intermittent and continuous auditory displays. *Proceedings of the 12th International Conference on Auditory Display (ICAD06)*, London, UK.
- Wickens, C. D., Gordon, S. E., & Liu, Y. (1998). *An introduction to human factors engineering*. New York: Longman.
- Wickens, C. D., & Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human Factors*, 30(5), 599-616.
- Wickens, C. D., Sandry, D. L., & Vidulich, M. (1983). Compatibility and resource competition between modalities of input, central processing, and output. *Human Factors*, 25(2), 227-248.
- Williams, T. M., & Aiken, L. S. (1975). Auditory pattern classification: Continuity of prototype use with development. *Developmental Psychology*, 11(6), 715-723.

- Williams, T. M., & Aiken, L. S. (1977). Development of pattern classification: Auditory-visual equivalence in the use of prototypes. *Developmental Psychology*, 13(3), 198-204.
- Zacks, J., Levy, E., Tversky, B., & Schiano, D. (2002). Graphs in print. In M. Anderson, B. Meyer & P. Olivier (Eds.), *Diagrammatic Representation and Reasoning* (pp. 187-206). London: Springer.