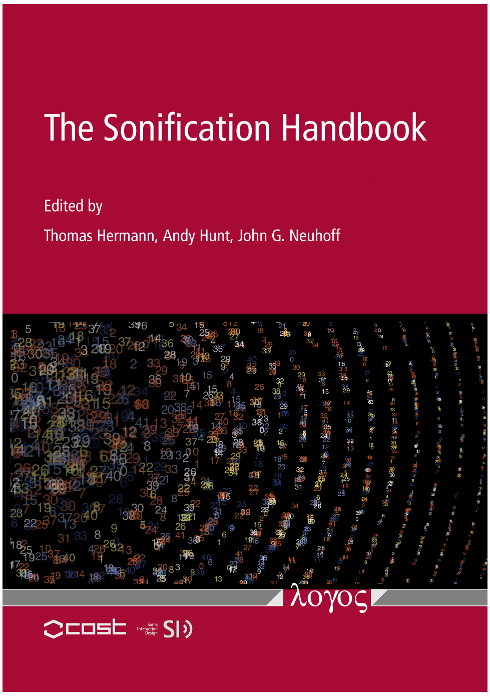

The Sonification Handbook

Edited by
Thomas Hermann, Andy Hunt, John G. Neuhoff

Logos Publishing House, Berlin, Germany
ISBN 978-3-8325-2819-5
2011, 586 pages

Online: http://sonification.de/handbook
Order: http://www.logos-verlag.com

Reference:
Hermann, T., Hunt, A., Neuhoff, J. G., editors (2011). *The Sonification Handbook*. Logos Publishing House, Berlin, Germany.

Chapter 2

Theory of Sonification

Bruce N. Walker and Michael A. Nees

This chapter give a broad introduction to sonification, and discusses the guiding theoretical considerations for sonification researchers and designers. It brings in many of the insights from relevant domains of research, and offers areas where future researchers could answer unresolved questions or make fruitful clarifications or qualifications to the field. While the chapter stands on its own as an overview of sonification, in many cases the interested reader is pointed to another more detailed chapter in this book, or to other external sources for more extensive coverage.

Reference:
Walker, B. N. and Nees, M. A. (2011). Theory of sonification. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook, chapter 2, pages 9–39. Logos Publishing House, Berlin, Germany.

Media examples: http://sonification.de/handbook/chapters/chapter2

Chapter 2

Theory of Sonification

Bruce N. Walker and Michael A. Nees

2.1 Chapter Overview

An auditory display can be broadly defined as any display that uses sound to communicate information. Sonification has been defined as a subtype of auditory displays that use non-speech 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", and this definition has persevered since its publication. More recently, a revised definition of sonification was proposed to both expand and constrain the definition of sonification to "..the data-dependent generation of sound, if the transformation is systematic, objective and reproducible..." (also see Hermann, 2008; Hermann, 2011). Sonification, then, seeks to translate relationships in data or information into sound(s) that exploit the auditory perceptual abilities of human beings such that the data relationships are comprehensible.

Theories offer empirically-substantiated, explanatory statements about relationships between variables. Hooker (2004) writes, "Theory represents our best efforts to make the world intelligible. It must not only tell us how things are, but why things are as they are" (pp. 74). Sonification involves elements of both science, which must be driven by theory, and design, which is not always scientific or theory-driven.

The theoretical underpinnings of research and design that can apply to (and drive) sonification come 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 sonification principles or rules (see Edworthy, 1998). Rather, the guiding principles of sonification in research and practice can be best characterized as an amalgam of important insights drawn from the convergence of these many diverse fields. While there have certainly been plenty of generalized contributions toward the sonification theory base (e.g., Barrass, 1997; Brazil, 2010; de Campo, 2007; Frauenberger & Stockman, 2009; Hermann, 2008; Nees & Walker, 2007; Neuhoff & Heller, 2005; Walker, 2002, 2007), to date, researchers and practitioners in sonification have yet to articulate a complete theoretical paradigm to guide research and design. Renewed interest and vigorous conversations on the topic have been reignited in recent years (see, e.g., Brazil & Fernstrom, 2009; de Campo, 2007; Frauenberger, Stockman, & Bourguet, 2007b; Nees & Walker, 2007).

The 1999 collaborative *Sonification Report* (Kramer et al., 1999) offered a starting point for a meaningful discussion of the theory of sonification by identifying 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. It attempts to draw upon the insights of relevant domains of research, and where necessary, offers areas where future researchers could answer unresolved questions or make fruitful clarifications or qualifications to the current state of the field. In many cases, the interested reader is pointed 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). In the case of an auditory display, the data of interest are conveyed to the human listener through sound.

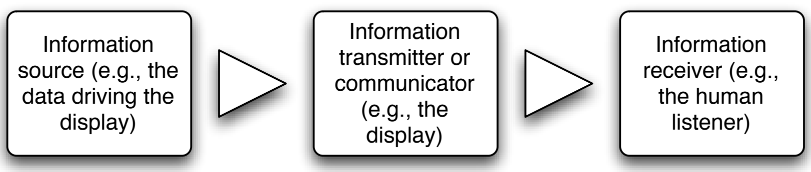

Figure 2.1: General description of a communication system.

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 inevitable given the ease and cost efficiency with which electronic devices 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[1] to deliver messages to the user. Despite these advances, there remains lingering doubt for some about the usefulness of sound in systems and ongoing confusion for many about how to implement sound in user interfaces (Frauenberger, Stockman, & Bourguet, 2007a).

The rationales and motivations for displaying information using sound (rather than a visual presentation, etc.) have been discussed extensively in the literature (e.g., Buxton et al., 1985; Hereford & Winn, 1994; Kramer, 1994; Nees & Walker, 2009; Peres et al., 2008; Sanderson, 2006). Briefly, though, auditory displays exploit the superior ability of the human 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.

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 decreasing in size, sound may be a compelling display mode as visual displays shrink (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 sonification 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

[1] *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.

and sonifications according to both function and technique, although, as our discussion will elaborate, they are very much inter-related.

Sonification is clearly a subset of auditory display, but it is not clear, in the end, where the exact boundaries should be drawn. Recent work by Hermann (2008) identified data-dependency, objectivity, systematicness, and reproducibility as the necessary and sufficient conditions for a sound to be called “sonification”. Categorical definitions within the sonification field, however, 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 different from each other in both form and function. Recent work on auditory displays that use speech-like sounds (Jeon & Walker, 2011; Walker, Nance, & Lindsay, 2006b) has even called into question the viability of excluding speech sounds from taxonomies of sonification (for a discussion, also see Worrall, 2009a).

Despite the difficulties with describing categories of auditory displays, such catalogs of auditory interfaces can be helpful to the extent that they standardize terminology and give the reader an idea of the options available for using sound in interfaces. In the interest of presenting a basic overview, this chapter provides a description, with definitions where appropriate, of the types of sounds that typically have been used in auditory interfaces. Other taxonomies and descriptions of auditory displays are available elsewhere (Buxton, 1989; de Campo, 2007; Hermann, 2008; Kramer, 1994; Nees & Walker, 2009), and a very extensive set of definitions for auditory displays (Letowski et al., 2001) has been published. 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 and the auditory display community. However, these descriptions should not be taken to imply that clear-cut boundaries and distinctions are always possible to draw or agree upon, 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, entertainment, sports, and exercise.

The following sections expand each of the above categories.

Alerting functions

Alerts and notifications refer to sounds used to indicate that something has occurred, 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 and chapter 19 in this volume). Warning signals presented in the auditory modality 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, Liu, & Jenkins, 2009) 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. Many (ten or more) alarms might be used in a single environment (Edworthy & Hellier, 2000), and Edworthy (2005) has critiqued the overabundance of alarms as a potential obstacle to the success of auditory alarms. Recent work (Edworthy & Hellier, 2006; Sanderson, 2006; Sanderson et al., 2009) has examined issues surrounding false alarms and suggested potential emerging solutions to reduce false alarms, including the design of intelligent systems that use multivariate input to look for multiple cues and redundant evidence of a real critical event. Sanderson et al. argued that the continuous nature of many sonifications effectively eliminates the problem of choosing a threshold for triggering a single discrete auditory warning. While it is clear that the interruptive and preemptive nature of sound is especially problematic for false alarms, more research is needed to understand whether sonifications or continuous auditory displays will alleviate this problem.

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* (also see chapter 18 in this volume). 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). Recent work (e.g., Jeon, Davison, Nees, Wilson, & Walker, 2009; Jeon & Walker, 2011; Walker, Nance, & Lindsay, 2006b) has begun to examine speech-like sounds for indicating a user's progress while scrolling auditory representations of common menu structures in devices (see sound examples **S2.1** and **S2.2**).

Data exploration functions

The third functional class of auditory displays contains those designed to permit *data exploration* (also see chapter 8 and 20 in this volume). 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, though some auditory displays, such as soundscapes (Mauney & Walker, 2004), blend status indicator and data exploration functions. *Auditory graphs* (for representative work, see Brown & Brewster, 2003; Flowers & Hauer, 1992, 1993, 1995; Smith & Walker, 2005) and model-based sonifications (see Chapter 11 in this volume and Hermann & Hunt, 2005) are typical exemplars of sonifications designed for data exploration purposes.

Entertainment, sports, and leisure

Auditory interfaces have been prototyped and researched in the service of exhibitions as well as leisure and fitness activities. Audio-only versions have appeared for simple, traditional games such as the Towers of Hanoi (Winberg & Hellstrom, 2001) and Tic-Tac-Toe (Targett & Fernstrom, 2003), and more complex game genres such as arcade games (e.g., space invaders, see McCrindle & Symons, 2000) and role-playing games (Liljedahl, Papworth, & Lindberg, 2007) have begun to appear in auditory-only formats.

Auditory displays also have been used to facilitate the participation of visually-impaired children and adults in team sports. Stockman (2007) designed an audio-only computer soccer game that may facilitate live action collaborative play between blind and sighted players. Sonifications have recently shown benefits as real-time biofeedback displays for competitive sports such as rowing (Schaffert, Mattes, Barrass, & Effenberg, 2009) and speed skating (Godbout & Boyd, 2010). While research in this domain has barely scratched the surface of potential uses of sonification for exercise, there is clearly a potential for auditory displays to give useful feedback and perhaps even offer corrective measures for technique (e.g., Godbout) in a variety of recreational and competitive sports and exercises (also see chapter 21 in this volume).

Auditory displays have recently been explored as a means of bringing some of the experience and excitement of dynamic exhibits to the visually impaired. A system for using sonified soundscapes to convey dynamic movement of fish in an "accessible aquarium" has been developed (Walker, Godfrey, Orlosky, Bruce, & Sanford, 2006a; Walker, Kim, & Pendse, 2007). Computer vision and other sensing technologies track the movements of entities within the exhibit, and these movements are translated, in real time, to musical representations. For example, different fish might be represented by different instruments. The location of an individual fish might be represented with spatialization of the sound while speed of movement is displayed with tempo changes. Soundscapes in dynamic exhibits may not only make such experiences accessible for the visually impaired, but may also enhance the experience for sighted viewers. Research (Storms & Zyda, 2000) has shown, for example, that high quality audio increases the perceived quality of concurrent visual displays in virtual environments. More research is needed to determine whether high quality auditory displays in dynamic exhibits enhance the perceived quality as compared to the visual experience alone.

Art

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. As one example, Quinn (2001, 2003) has used data sonifications to drive ambitious musical works, and he has produced entire albums of compositions. Of note, the mapping of data to sound must be systematic in compositions, and the potentially subtle distinction between sonification and music as a conveyor of information is debatable (see Worrall, 2009a). Vickers and Hogg (2006) offered a seminal discussion of the similarities between sonification and music.

2.3.2 Sonification techniques and approaches

Another way to organize and define sonifications is to describe them according to their sonification technique or approach. de Campo (2007) offered a sonification design map (see Figure 10.1 on page 252) that featured three broad categorizations of sonification approaches:

1. event-based;
2. model-based; and
3. continuous.

de Campo's (2007) approach is useful in that it places most non-speech auditory displays within a design framework. The appeal of de Campo's approach is its placement of different types of auditory interfaces along continua that allow for blurry boundaries between categories, and the framework also offers some guidance for choosing a sonification technique. Again, the definitional boundaries to taxonomic descriptions of sonifications are indistinct and often overlapping. Next, a brief overview of approaches and techniques employed in sonification is provided; but for a more detailed treatment, see Part III of this volume.

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 (also see chapter 11 in this volume). 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 Brown, Brewster, & Riedel, 2002). Sonifications more toward this interactive end of the spectrum have been called "conversation mode" (Walker & Kramer, 1996) or "query based" (Franklin & Roberts, 2004) sonification. In other cases, user input and interaction may be the required catalyst that drives the presentation of sounds (see 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).

Parameter mapping sonification

Parameter mapping represents changes in some data dimension with changes in an acoustic dimension to produce a sonification (see chapter 15 in this volume). 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 lower dimension display than the model-based approaches discussed below. The data changes may be more qualitative or discrete, such as a thresholded on or off response that triggers a discrete alarm, or parameter mapping may be used with a series of discrete data points to produce a display that seems more continuous. These 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.

Model-based sonification

Model-based approaches to sonification (Hermann, 2002, chapter 16 in this volume; Hermann & Ritter, 1999) differ from event-based approaches in that instead of mapping data parameters to sound parameters, the display designer builds a virtual model whose sonic responses to user input are derived from data. A model, then, is a virtual object or instrument with which the user can interact, and the user's input drives the sonification such that "the sonification is the reaction of the data-driven model to the actions of the user" (Hermann, 2002 p. 40). 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. Model-based approaches rely upon (and the sounds produced are contingent upon) the active manipulation of the sonification by the user. These types of sonifications tend to involve high data dimensionality and large numbers of data points.

Audification

Audification is the most prototypical method of direct sonification, whereby waveforms of periodic data are directly translated into sound (Kramer, 1994, chapter 12 in this volume). 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.

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 ones used for alerts, notifications, alarms, and even status and process monitors, as these functions are all triggered by events in the system being monitored. Data exploration may employ event-based approaches, model-based sonification, or continuous sonification depending upon the specific task of the user (Barrass, 1997).

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[2]) are relevant to the user's task;

[2]The terms "data" and "information" are used more or less interchangeably here in a manner consistent with Hermann's (2008) definition of sonification. For other perspectives, see Barrass (1997) or Worrall (2009b, Chapter 3)

- 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. Mapping data to sound requires a consideration of perceptual or "bottom up" processes, 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). Another challenge comes from the more cognitive or conceptual "top down" components of perceiving sonifications. 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). 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 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 Stevens, 1975; Walker, 2007 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[3]. A discussion of the types of tasks that users might undertake with sonifications, therefore, closely parallels the taxonomies of auditory displays described above.

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 crossing 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) 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.

Awareness of a process or situation

Sonifications may sometimes be employed to promote the awareness of task-related processes or situations (also see chapter 18 in this volume). Awareness-related task goals are different

[3] 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) or Barrass (1996) for thorough coverage of task analysis purposes and methods.

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 task-related processes.

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

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.

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.

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 narrower, 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.

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 data and complex relationships among multiple variables. Through interactions with the virtual object, the listener hopes to extract information about the relationships within, and structure of, the data represented.

Exploratory inspection Occasionally, a user's task may be entirely exploratory requiring the inspection or examination of data with no a priori questions in mind. Kramer (1994) described exploratory tasks with sound as a less tractable endeavor 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.

Dual task performance and multimodal tasking scenarios

In many applications of sonification, 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. Bonebright and Nees (2009) presented sounds that required a manual response approximately every 6 seconds while participants listened to a passage for verbal comprehension read aloud. The sound used included five types of earcons and also brief speech sounds, and the researchers predicted that speech sounds would interfere most with spoken passage comprehension. Surprisingly, however, only one condition—featuring particularly poorly designed earcons that used a continuous pitch-change mapping—significantly interfered with passage comprehension compared to a control condition involving listening only without the concurrent sound task. Although speech sounds and the spoken passage presumably taxed the same verbal working memory resources, and all stimuli were concurrently delivered to the ears, there was little dual-task effect, presumably because the sound task was not especially hard for participants.

Despite these encouraging results, a wealth of questions abounds 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 non-speech 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

Semiotics is "the science of signs (and signals)" (Cuddon, 1991 p. 853). Clearly sonification aims to use sound to signify data or other information (Barrass, 1997), 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) and is presented to a design panel of experts or representative users. Such an approach seeks input regarding the most appropriate way to use sounds as signs 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 (e.g., the sound of branches snapping to represent a tornado). Finally, indirect metaphorical relationships are those in which the sound signal is related to its referent only in some emblematic way (e.g., the sound of a mosquito buzzing to represent a helicopter).

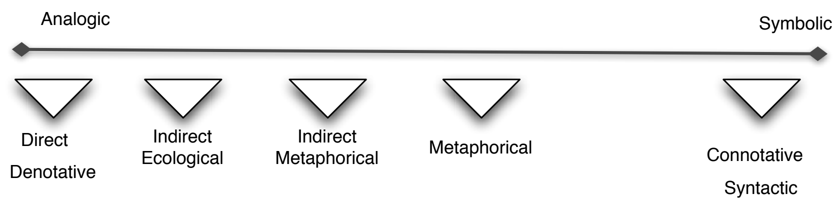

Figure 2.2: The analogic-symbolic representation continuum.

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 (see chapter 13 in this volume). In other words, the sound bears some ecological (i.e., naturally-associated) resemblance to the action or process (see Gaver, 1994; Kramer, 1994). This approach has also been called *nomic 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 and chapter 14 in this volume). 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 tradeoff exists in that the abstract nature of earcons may require longer learning time or even formal training in their use. Walker and colleagues (Palladino & Walker, 2007; Walker & Kogan, 2009; Walker et al., 2006b) have discussed a new type of interface sound, the *spearcon*, which is intended to overcome the shortcomings of both auditory icons and earcons. Spearcons (see sound examples **S2.1** and **S2.3**) are created by speeding up a spoken phrase even 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, 2007; Walker & Kramer, 2004).

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 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, 2007). 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; Brown, Brewster, Ramloll, Burton, & Riedel, 2003; Flowers & Hauer, 1995; Neuhoff & Heller, 2005; Smith & Walker, 2005; Walker, 2002, 2007).

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 (Mauney & Walker, 2010; Walker & Lane, 2001). Walker (2002, 2007) 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.

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 (for early work on scaling a perceptual sound space, see Barrass, 1994/2005). This is a key distinction between sonifications and warnings or trend monitoring sounds. Again, Walker (2002, 2007) 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, 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.

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) and spatial location (or stereo panning) are parameters that sonification designers can often use simply and effectively (see also Bonebright et al., 2001; 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.

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 without context cues.

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 (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 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 (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 (e.g., musical, etc.) sonifications to the extent possible while still conveying the intended message. Vickers and Hogg (2006) made a pointed statement about aesthetics in sonification. In particular, they argued that more careful attention to aesthetics would facilitate ease of listening with sonifications, which would in turn promote comprehension of the intended message of the displays.

2.6.2 Individual differences and training

The capabilities, limitations, and experiences of listeners, as well as transient states (such as 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 most 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.

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 chapter 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; 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 (also see Neuhoff, 2004). 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 judgments of discrimination (e.g., just-noticeable differences, etc.) and identification of sounds. Again, however, the data regarding discrimination or identification performance in controlled conditions may offer misleading design heuristics for less controlled, non-laboratory environments. Sonification researchers can and should, however, actively borrow from and adapt the knowledge and methods of psychoacousticians. For example, Bregman's (1990) theory of auditory scene analysis (ASA) has considerable explanatory value with respect to the pre-attentive emergence of auditory objects and gestalts, and this perspective can offer auditory display design heuristics (see, e.g., Barrass & Best, 2008). Similarly, Sandor and Lane (2003) introduced the term *mappable difference* to describe the absolute error in response accuracy 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.

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. 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. A recent qualitative study (Nees & Walker, 2008) and a series of experiments (Nees, 2009; Nees & Walker, in press) have both suggested that non-speech sound can be rehearsed in working memory as words, visual images, or as quasi-isomorphic sounds per se. 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 favorably inform display design.

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., 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 non-musicians have been reported (e.g., Lacherez, Seah, & Sanderson, 2007; 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 are likely have auditory perceptual abilities on par with the best musicians.

Visually-impaired versus sighted listeners

Though sonification research is most often accomplished with samples of sighted students in academic settings, auditory displays may provide enhanced accessibility to information for visually-impaired listeners. Visual impairment represents an individual difference that has been shown to have a potentially profound impact on the perception of sounds in some scenarios. Walker and Lane (2001), for example, showed that blind and sighted listeners actually had opposing intuitions about the polarity of the pairing of some acoustic dimensions with conceptual data dimensions. Specifically, blind listeners expected that increasing frequency represented a decreasing “number of dollars” (a negative polarity) whereas sighted listeners expected that increasing frequency conveyed that wealth was accumulating (a positive polarity). This finding was extended upon and further confirmed in a recent study (Mauney & Walker, 2010). These data also suggested that, despite generally similar patterns of magnitude estimation for conceptual data dimensions, sighted participants were more likely to intuit split polarities than blind participants. Individual differences between visually-impaired and sighted listeners require more research and a careful testing of auditory displays with the intended user population. Potential differences between these user groups are not necessarily predictable from available design heuristics.

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.

2.7 Conclusions: Toward a Cohesive Theoretical Account of Sonification

Current research is taking the field of sonification in many exciting directions, and researchers and practitioners have only just begun to harness the potential for sound to enhance and improve existing interfaces or be developed into purely auditory interfaces. The literature on auditory displays has grown tremendously. These successes notwithstanding, sonification research and design faces many obstacles and challenges in the pursuit of ubiquitous, usable, and aesthetically pleasing sounds for human-machine interactions, and perhaps the most pressing obstacle is the need for a cohesive theoretical paradigm in which research and design can continue to develop.

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 account of sound as an information display medium. To date, few theories or models of human interaction with auditory displays exist. It seems inevitable that the field of sonification will need to develop fuller explanatory models in order to realize the full potential of the field. As Edwards (1989) 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. These benefits notwithstanding, the development of theory remains difficult, especially in pragmatic and somewhat design-oriented fields like sonification (for a discussion, see Hooker, 2004).

A distinction has been drawn, however, between "theorizing" as a growing process within a field, and "theory" as a product of that process (Weick, 1995). Despite the absence of a grand theory of sonification, recent developments reflect the field's active march toward meaningful theory. Important evidence of progress toward meeting some of the conditions of a cohesive theory of sonification is emerging. Theory in sonification will depend upon a shared language, and Hermann (2008) recently initiated a much-needed discussion about definitional boundaries and fundamental terminology in the field. Theory requires a meaningful organization of extant knowledge, and de Campo's (2007) recent work offered an important step toward describing the diverse array of sonification designs within a common space. Theory will bridge the gap between research and practice, and Brazil (Brazil, 2010;

Brazil & Fernstrom, 2009) has begun to offer insights for integrating sonification design and empirical methods of evaluation (also see chapter 6 in this volume). Theory specifies the important variables that contribute to performance of the data-display-human system. Nees and Walker (2007) recently described a conceptual model of the variables relevant to auditory graph comprehension, whereas Bruce and Walker (2009) took a similar conceptual model approach toward understanding the role of audio in dynamic exhibits. Theory will result in reusable knowledge rather than idiosyncratic, ad hoc designs, and Frauenberger and Stockman (2009) have developed a framework to assist in the capture and dissemination of effective designs for auditory displays. As such, there is reason for optimism about the future of theoretical work in the field of sonification, and a shared based of organized knowledge that guides new research and best practice implementation of sonifications should be one of the foremost aspirations of the field in the immediate future.

Bibliography

[1] Barrass, S. (1994/2005). A perceptual framework for the auditory display of scientific data. *ACM Transactions on Applied Perception*, 2(4), 389–402.

[2] Barrass, S. (1996). TaDa! Demonstrations of auditory information design. *Proceedings of the 3rd International Conference on Auditory Display*, Palo Alto, CA.

[3] Barrass, S. (1997). Auditory information design. Unpublished Dissertation, Australian National University.

[4] Barrass, S. (2005). A perceptual framework for the auditory display of scientific data. *ACM Transactions on Applied Perception*, 2(4), 389–492.

[5] Barrass, S., & Best, V. (2008). Stream-based sonification diagrams. *Proceedings of the 14th International Conference on Auditory Display*, Paris, France.

[6] Bertin, J. (1983). *Semiology of Graphics* (W. J. Berg, Trans.). Madison, Wisconsin: The University of Wisconsin Press.

[7] 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.

[8] Bonebright, T. L., & Nees, M. A. (2009). Most earcons do not interfere with spoken passage comprehension. *Applied Cognitive Psychology*, 23(3), 431–445.

[9] 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.

[10] Brazil, E. (2010). A review of methods and frameworks for sonic interaction design: Exploring existing approaches. *Lecture Notes in Computer Science*, 5954, 41–67.

[11] Brazil, E., & Fernstrom, M. (2009). Empirically based auditory display design. *Proceedings of the SMC 2009 – 6th Sound and Computing Conference* (pp. 7–12), Porto, Portugal.

[12] Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound.* Cambridge, MA: MIT Press.

[13] Brewster, S. (1997). Using non-speech sound to overcome information overload. *Displays*, 17, 179–189.

[14] Brewster, S., & Murray, R. (2000). Presenting dynamic information on mobile computers. *Personal Technologies*, 4(4), 209–212.

[15] Brown, A. W. (1928). The reliability and validity of the Seashore Tests of Musical Talent. *Journal of Applied Psychology*, 12, 468–476.

[16] Brown, L. M., Brewster, S., & Riedel, B. (2002). Browsing modes for exploring sonified line graphs. *Proceedings of the 16th British HCI Conference*, London, UK.

[17] 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.

[18] 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.

[19] 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).

[20] Bruce, C., & Walker, B. N. (2009). Modeling visitor-exhibit interaction at dynamic zoo and aquarium exhibits for developing real-time interpretation. *Proceedings of the Association for the Advancement of Assistive Technology in Europe Conference* (pp. 682–687).

[21] Buxton, W. (1989). Introduction to this special issue on nonspeech audio. *Human-Computer Interaction*, 4, 1–9.

[22] Buxton, W., Bly, S. A., Frysinger, S. P., Lunney, D., Mansur, D. L., Mezrich, J. J., et al. (1985). Communicating with sound. *Proceedings of the CHI '85* (pp. 115–119).

[23] Cary, H. (1923). Are you a musician? Professor Seashore's specific psychological tests for specific musical abilities. *Scientific American*, 326–327.

[24] 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.

[25] 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.

[26] Cuddon, J. A. (1991). *Dictionary of Literary Terms and Literary Theory* (3rd ed.). New York: Penguin Books.

[27] Cusack, R., & Roberts, B. (2000). Effects of differences in timbre on sequential grouping. *Perception & Psychophysics*, 62(5), 1112–1120.

[28] de Campo, A. (2007). Toward a data sonification design space map. *Proceedings of the International Conference on Auditory Display (ICAD2007)* (pp. 342–347), Montreal, Canada.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] Edworthy, J., & Hellier, E. (2000). Auditory warnings in noisy environments. *Noise & Health*, 2(6), 27–39.

[34] Edworthy, J., & Hellier, E. (2005). Fewer but better auditory alarms will improve patient safety. *British Medical Journal*, 14(3), 212–215.

[35] 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.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] 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.

[41] 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.

[42] 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).

[43] Frauenberger, C., & Stockman, T. (2009). Auditory display design: An investigation of a design pattern approach. *International Journal of Human-Computer Studies*, 67, 907–922.

[44] Frauenberger, C., Stockman, T., & Bourguet, M.-L. (2007). A Survey on Common Practice in Designing Audio User Interface. *21st British HCI Group Annual Conference (HCI 2007).* Lancaster, UK.

[45] Frauenberger, C., Stockman, T., & Bourguet, M.-L. (2007). Pattern design in the context space: A methodological framework for auditory display design. *Proceedings of the International Conference on Auditory Display (ICAD2007)* (pp. 513–518), Montreal, Canada.

[46] 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.

[47] Frysinger, 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.

[48] 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.

[49] Garner, W. R., & Gottwald, R. L. (1968). The perception and learning of temporal patterns. *The Quarterly Journal of Experimental Psychology*, 20(2).

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

[51] Gaver, W. W. (1993). What in the world do we hear? An ecological approach to auditory event perception. *Ecological Psychoogy*, 5(1), 1–29.

[52] 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.

[53] 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.

[54] Godbout, A., & Boyd, J. E. (2010). Corrective sonic feedback for speed skating: A case study. *Proceedings of the 16th International Conference on Auditory Display (ICAD2010)* (pp. 23–30), Washington, DC.

[55] 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

[56] Hereford, J., & Winn, W. (1994). Non-speech sound in human-computer interaction: A review and design guidelines. *Journal of Educational Computer Research*, 11, 211–233.

[57] Hermann, T. (2002). Sonification for exploratory data analysis. Ph.D. thesis, Faculty of Technology, Bielefeld University, http://sonification.de/publications/Hermann2002-SFE

[58] Hermann, T. (2008). Taxonomy and definitions for sonification and auditory display. *Proceedings of the 14th International Conference on Auditory Display*, Paris, France.

[59] Hermann, T. (2011). Sonification – A Definition. Retrieved January 23, 2011, from http://sonification.de/son/definition

[60] Hermann, T., & Hunt, A. (2005). An introduction to interactive sonification. *IEEE Multimedia*, 12(2), 20–24.

[61] Hermann, T., & Ritter, H. (1999). Listen to your data: Model-based sonification for data analysis. In G. D. Lasker (Ed.), *Advances in Intelligent Computing and Multimedia Systems* (pp. 189–194). Baden-Baden, Germany: IIASSRC.

[62] Hooker, J. N. (2004). Is design theory possible? *Journal of Information Technology Theory and Application*, 6(2), 73–82.

[63] 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.

[64] 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.

[65] Jeon, M., Davison, B., Nees, M. A., Wilson, J., & Walker, B. N. (2009). Enhanced auditory menu cues improve dual task performance and are preferred with in-vehicle technologies. *Proceedings of the First International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive UI 2009)*, Essen, Germany.

[66] Jeon, M., & Walker, B. N. (2011). Spindex (speech index) improves auditory menu acceptance and navigation performance. *ACM Transactions on Accessible Computing, 3,* Article 10.

[67] Johannsen, G. (2004). Auditory displays in human-machine interfaces. *Proceedings of the IEEE*, 92(4), 742–758.

[68] 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.

[69] 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.

[70] 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.

[71] 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).

[72] Lacherez, P., Seah, E. L., & Sanderson, P. M. (2007). Overlapping melodic alarms are almost indiscriminable. *Human Factors*, 49(4), 637–645.

[73] Lakatos, S., McAdams, S., & Causse, R. (1997). The representation of auditory source characteristics: Simple geometric form. *Perception & Psychophysics*, 59(8), 1180–1190.

[74] Letowski, T., Karsh, R., Vause, N., Shilling, R. D., Ballas, J., Brungart, D., et al. (2001). Human factors military lexicon: Auditory displays. Army Research Laboratory Technical Report No. ARL-TR-2526

[75] 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.

[76] Liljedahl, M., Papworth, N., & Lindberg, S. (2007). Beowulf: A game experience built on sound effects. *Proceedings of the International Conference on Auditory Display (ICAD2007)* (pp. 102–106), Montreal, Canada.

[77] Luczak, H. (1997). Task analysis. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (2nd ed., pp. 340–416). New York: Wiley.

[78] 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.

[79] Mauney, L. M., & Walker, B. N. (2010). Universal design of auditory graphs: A comparison of sonification mappings for visually impaired and sighted listeners. *ACM Transactions on Accessible Computing*, 2(3), Article 12.

[80] McAdams, S., & Bigand, E. (1993). *Thinking in Sound: The Cognitive Psychology of Human Audition.*

Oxford: Oxford University Press.

[81] McCrindle, R. J., & Symons, D. (2000). Audio space invaders. *Proceedings of the 3rd International Conference on Disability, Virtual Reality, & Associated Technologies* (pp. 59–65), Alghero, Italy.

[82] McGookin, D. K., & Brewster, S. A. (2004). Understanding concurrent earcons: Applying auditory scene analysis principles to concurrent earcon recognition. *ACM Transactions on Applied Perception*, 1, 130–150.

[83] Meyer, J. (2000). Performance with tables and graphs: Effects of training and a visual search model. *Ergonomics*, 43(11), 1840–1865.

[84] Meyer, J., Shinar, D., & Leiser, D. (1997). Multiple factors that determine performance with tables and graphs. *Human Factors*, 39(2), 268–286.

[85] Moore, B. C. J. (1997). *An Introduction to the Psychology of Hearing* (4th ed.). San Diego, Calif.: Academic Press.

[86] Mulligan, B. E., McBride, D. K., & Goodman, L. S. (1984). A design guide for nonspeech auditory displays: Naval Aerospace Medical Research Laboratory Technical Report, Special Report No. 84–1.

[87] Nees, M. A. (2009). Internal representations of auditory frequency: Behavioral studies of format and malleability by instructions. Unpublished Ph.D. Dissertation.: Georgia Institute of Technology. Atlanta, GA.

[88] 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.

[89] Nees, M. A., & Walker, B. N. (2007). Listener, task, and auditory graph: Toward a conceptual model of auditory graph comprehension. *Proceedings of the International Conference on Auditory Display (ICAD2007)* (pp. 266–273), Montreal, Canada.

[90] Nees, M. A., & Walker, B. N. (2008). Encoding and representation of information in auditory graphs: Descriptive reports of listener strategies for understanding data. *Proceedings of the International Conference on Auditory Display (ICAD 08)*, Paris, FR (24–27 June).

[91] Nees, M. A., & Walker, B. N. (2009). Auditory interfaces and sonification. In C. Stephanidis (Ed.), *The Universal Access Handbook* (pp. TBD). New York: CRC Press.

[92] Nees, M. A., & Walker, B. N. (in press). Mental scanning of sonifications reveals flexible encoding of nonspeech sounds and a universal per-item scanning cost. *Acta Psychologica*.

[93] Neuhoff, J. G. (Ed.). (2004). *Ecological Psychoacoustics*. New York: Academic Press.

[94] 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.

[95] 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.

[96] 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.

[97] Palladino, D., & Walker, B. N. (2007). Learning rates for auditory menus enhanced with spearcons versus earcons. *Proceedings of the International Conference on Auditory Display (ICAD2007)* (pp. 274–279), Montreal, Canada.

[98] Peres, S. C., Best, V., Brock, D., Shinn-Cunningham, B., Frauenberger, C., Hermann, T., et al. (2008). Auditory Interfaces. In P. Kortum (Ed.), *HCI Beyond the GUI: Design for Haptic, Speech, Olfactory and Other Nontraditional Interfaces* (pp. 147–196). Burlington, MA: Morgan Kaufmann.

[99] 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.

[100] 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.

[101] 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.

[102] Quinn, M. (2003). For those who died: A 9/11 tribute. *Proceedings of the 9th International Conference on Auditory Display*, Boston, MA.

[103] Salvendy, G. (1997). *Handbook of Human Factors and Ergonomics* (2nd ed.). New York: Wiley.

[104] Sanders, M. S., & McCormick, E. J. (1993). *Human Factors in Engineering and Design* (7th ed.). New York: McGraw-Hill.

[105] Sanderson, P. M. (2006). The multimodal world of medical monitoring displays. *Applied Ergonomics*, 37, 501–512.

[106] Sanderson, P. M., Liu, D., & Jenkins, D. A. (2009). Auditory displays in anesthesiology. *Current Opinion in Anesthesiology*, 22, 788–795.

[107] 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.

[108] Schaffert, N., Mattes, K., Barrass, S., & Effenberg, A. O. (2009). Exploring function and aesthetics in sonifications for elite sports. *Proceedings of the Second International Conference on Music Communication Science*, Sydney, Australia.

[109] Seashore, C. E., Lewis, D., & Saetveit, J. G. (1960). *Seashore Measures of Musical Talents* (Revised 1960 ed.). New York: The Psychological Corp.

[110] 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.

[111] Smith, D. R., & Walker, B. N. (2005). Effects of auditory context cues and training on performance of a point estimation sonification task. *Applied Cognitive Psychology*, 19(8), 1065–1087.

[112] 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.

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

[114] 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.

[115] Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 13(2684), 677–680.

[116] Stevens, S. S. (1975). *Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects*. New York: Wiley.

[117] Stockman, T., Rajgor, N., Metatla, O., & Harrar, L. (2007). The design of interactive audio soccer. *Proceedings of the 13th International Conference on Auditory Display* (pp. 526–529), Montreal, Canada.

[118] Stokes, A., Wickens, C. D., & Kite, K. (1990). *Display Technology: Human Factors Concepts*. Warrendale, PA: Society of Automotive Engineers.

[119] Storms, R. L., & Zyda, M. J. (2000). Interactions in perceived quality of auditory-visual displays. *Presence: Teleoperators & Virtual Environments*, 9(6), 557–580.

[120] Targett, S., & Fernstrom, M. (2003). Audio games: Fun for all? All for fun? *Proceedings of the International Conference on Auditory Display (ICAD2003)* (pp. 216–219), Boston, MA.

[121] 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.

[122] 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.

[123] Tufte, E. R. (1990). *Envisioning Information*. Cheshire, Connecticut: Graphics Press.

[124] Vickers, P., & Hogg, B. (2006). Sonification abstraite/sonification concrete: An 'aesthetic perspective space' for classifying auditory displays in the ars musica domain. *Proceedings of the International Conference on Auditory Display (ICAD2006)* (pp. 210–216), London, UK.

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

[126] Walker, B. N. (2007). Consistency of magnitude estimations with conceptual data dimensions used for sonification. *Applied Cognitive Psychology*, 21, 579–599.

[127] 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.

[128] Walker, B. N., Godfrey, M. T., Orlosky, J. E., Bruce, C., & Sanford, J. (2006a). Aquarium sonification: Soundscapes for accessible dynamic informal learning environments. *Proceedings of the International Conference on Auditory Display (ICAD 2006)* (pp. 238–241), London, UK.

[129] Walker, B. N., Kim, J., & Pendse, A. (2007). Musical soundscapes for an accessible aquarium: Bringing dynamic exhibits to the visually impaired. *Proceedings of the International Computer Music Conference (ICMC 2007)* (pp. TBD), Copenhagen, Denmark.

[130] Walker, B. N., & Kogan, A. (2009). Spearcons enhance performance and preference for auditory menus on a mobile phone. *Proceedings of the 5th international conference on universal access in Human-Computer Interaction (UAHCI) at HCI International 2009*, San Diego, CA, USA.

[131] 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.

[132] 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.

[133] Walker, B. N., & Kramer, G. (2005). Mappings and metaphors in auditory displays: An experimental assessment. *ACM Transactions on Applied Perception*, 2(4), 407–412.

[134] 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.

[135] 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.

[136] 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.

[137] Walker, B. N., Nance, A., & Lindsay, J. (2006b). Spearcons: Speech-based earcons improve navigation performance in auditory menus. *Proceedings of the International Conference on Auditory Display (ICAD06)* (pp. 63–68), London, UK.

[138] Walker, B. N., & Nees, M. A. (2005). An agenda for research and development of multimodal graphs. *Proceedings of the International Conference on Auditory Display (ICAD2005)* (pp. 428–432), Limerick, Ireland.

[139] Walker, B. N., & Nees, M. A. (2005). Brief training for performance of a point estimation task sonification task. *Proceedings of the International Conference on Auditory Display (ICAD2005)*, Limerick, Ireland.

[140] 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.

[141] 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.

[142] Weick, K. E. (1995). What theory is not, theorizing is. *Administrative Science Quarterly*, 40(3), 385–390.

[143] Wickens, C. D., Gordon, S. E., & Liu, Y. (1998). *An Introduction to Human Factors Engineering*. New York:

Longman.

[144] Wickens, C. D., & Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human Factors*, 30(5), 599–616.

[145] 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.

[146] Winberg, F., & Hellstrom, S. O. (2001). Qualitative aspects of auditory direct manipulation: A case study of the Towers of Hanoi. *Proceedings of the International Conference on Auditory Display (ICAD 2001)* (pp. 16–20), Espoo, Finland.

[147] Worrall, D. R. (2009). An introduction to data sonification. In R. T. Dean (Ed.), *The Oxford Handbook of Computer Music and Digital Sound Culture* (pp. 312–333). Oxford: Oxford University Press.

[148] Worrall, D. R. (2009). Information sonification: Concepts, instruments, techniques. Unpublished Ph.D. Thesis: University of Canberra.

[149] 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.