Data Density and Trend Reversals in Auditory Graphs: Effects on Point-Estimation and Trend-Identification Tasks

MICHAEL A. NEES and BRUCE N. WALKER
Georgia Institute of Technology

Auditory graphs—displays that represent quantitative information with sound—have the potential to make data (and therefore science) more accessible for diverse user populations. No research to date, however, has systematically addressed the attributes of data that contribute to the complexity (the ease or difficulty of comprehension) of auditory graphs. A pair of studies examined the role of data density (i.e., the number of discrete data points presented per second) and the number of trend reversals for both point-estimation and trend-identification tasks with auditory graphs. For the point-estimation task, more trend reversals led to performance decrements. For the trend-identification task, a large main effect was again observed for trend reversals, but an interaction suggested that the effect of the number of trend reversals was different across lower data densities (i.e., as density increased from 1 to 2 data points per second). Results are discussed in terms of data sonification applications and rhythmic theories of auditory pattern perception.

Categories and Subject Descriptors: H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems—*Audio input/output*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Auditory (nonspeech) feedback*; H.5.5 [**Information Interfaces and Presentation**]: Sound and Music Computing—*Methodologies and techniques*

General Terms: Experimentation, Human factors

Additional Key Words and Phrases: Auditory graphs, sonification, auditory display

ACM Reference Format:

Nees, M. A. and Walker, B. N. 2008. Data density and trend reversals in auditory graphs: Effects on point-estimation and trend-identification tasks. ACM Trans. Appl. Percpt. 5, 3, Article 13 (August 2008), 24 pages. DOI = 10.1145/1402236.1402237 http://doi.acm.org/10.1145/1402236.1402237

1. INTRODUCTION

Graphical representations of quantitative information play an important role in the dissemination of information in both public and scientific forums. R.W. Jones and Carreras [1996] estimated that 2.2 trillion graphs were published via print media in 1994 alone. Zacks et al. [2002] sampled academic journals and newspapers across 10 years (1985–1994). They found that the average number of graphs published per issue in academic journals rose from 34.7 to 61.2 during the sampled time period, while the average number of graphs appearing in newsprint rose from 10.1 to 24.5 per issue. Not surprisingly, Peden and Hausmann [2000] found an average of 67.73 graphs per book in a sample of introductory psychology texts.

Portions of this research were supported by NSF Career Award #IIS-0644076 to Bruce N. Walker.
Authors' addresses: Michael A. Nees and Bruce N. Walker, Georgia Institute of Technology, School of Psychology, 654 Cherry Street, Atlanta, GA, 30332; email: Bruce.walker@psych.gatech.edu.

Graphical information displays pervade published material, because they can offer a relatively concise visual summary of data. Furthermore, properly designed graphs can facilitate the emergence of data features (e.g., patterns; see Sanderson et al. [1989]) that are not immediately evident in nongraphical data depictions; thus, researchers (e.g., Kosslyn [1989]) have theorized such data features in graphical displays are automatically perceived by the viewer. As a result of their prevalence and advantages over text for quantitative information display, graphs have been empirically studied and discussed both as instructional aids for learning in children or novice populations [Gobbo 1994; Lee and Gerber 1999; Liu et al. 1999; Moore 1993] and more generally as communicative information displays (see Butler [1993]; Carpenter and Shah [1998]; Carswell [1992]; Gillan and Lewis [1994]; Kosslyn [1989]; Meyer et al. [1997]). The representation of quantitative data, however, need not be confined to visual displays; scientists began to investigate the potential to represent data in the auditory modality over 50 years ago (for a brief review, see Frysinger [2005]). More recently, researchers have begun to examine auditory graphs as tools for data exploration and analysis.

1.1 Rationale: The Need for Auditory Graphs

Digital technology has allowed for the economical and widespread production and implementation of sounds via personal computers [Flowers et al. 2005]. The most obvious motivation for pursuing sound as a means of quantitative information display has been the inability of some individuals to access graphical visualizations (e.g., blind and visually impaired people). Kramer [1994], however, reviewed a number of users, tasks, and scenarios where auditory displays could be advantageous, and researchers have argued that auditory graphs have the potential to facilitate comprehension of graphical information for both blind and sighted students and scientists [Flowers et al. 2005; Walker and Nees 2005a]. Some environmental conditions or constraints of a system can render visual displays unusable (such as when a person's line of sight with the display is obscured [Kramer 1994]) or insufficient (like in cellular phones or other mobile devices where screen sizes are exceedingly small; see Brewster [2002]). On the other hand, technological and computing advances have also allowed for substantial *increases* in the size and number of visual displays in the typical computing workstation. This extra screen space can introduce problems, particularly with the overabundance and disorganization of visual information (for a discussion, see Grudin [2001]). Indeed, Wickens' [2002] description of multiple resources makes theoretical predictions that suggest less dual task interference when information can be spread across modalities rather than processed within a single modality (e.g., by vision alone). Studies have confirmed that auditory displays can be beneficial in scenarios where the display is small [Brewster 2002; Brewster and Murray 2000] or when the display is such that vision may be overtaxed [Brewster 1997; M. L. Brown et al. 1989].

For blind or visually impaired users of software and the world wide web, screen readers display information through a text-to-speech conversion. Images may be "tagged" with a hidden text description of the image, such as when a screen reader encounters a picture of a dog and reads aloud the tag "picture of dog." This commonly used and often recommended (e.g., Massachusetts Institute of Technology [n.d.]; Quesenbery [1999]; W3C [2000]) tagging strategy may make some media (i.e., pictures and photographs) suitably accessible for the blind. Graphs like those that pervade print media, however, are not easy to represent with a simple text tag. Furthermore, any attempt to verbally describe a graphical display of quantitative information results in a lengthy text description of the graph that may remove the advantages of the graphical display. To make the quantitative information "accessible," tables have often been used to represent data in lieu of more appropriate line graphs, a practice that also may remove the advantages of graphical representations (for some guidelines for presenting information in graphical form, see American Psychological Association [2001]; Gillan et al. [1998]; Oliver [1998]), Other current practices in graphing for the visually impaired include the use of tactile or Braille graphs, but the hardware to produce tangible graphics has been cost prohibitive. Auditory representations of quantitative information could fill a gap in data accessibility for the blind or help sighted people interact with computers and other digital technologies more efficiently.

1.2 What Is An Auditory Graph?

Auditory graphs are a class of displays that rely on sonification—nonspeech auditory information display [Kramer et al. 1999]. Auditory graphs most commonly map data changes along the visual y-axis to changes in the frequency of sounds, while visual x-axis changes are mapped to the temporal presentation of the sound (see, for example, Bonebright et al. [2001]; L. M. Brown and Brewster [2003]; L. M. Brown et al. [2003]; Childs [2005]; Flowers and Hauer [1995]; Mansur et al. [1985]; Smith and Walker [2002; 2005]; Walker and Nees [2005b]). Empirical evidence has suggested that changes in auditory frequency can be a perceptual analog for graphical spatial position, and theoretical approaches to the perception of pitch have described a dimension that corresponds to the "height" of frequency (see Shepard [1982]). Kubovy's Theory of Indispensable Attributes (TIA; see Kubovy [1981]) proposed that "visual spatial location is analogous to auditory frequency" (pp. 78). Frequency, then, is perhaps the most compelling dimension of sound for representations of quantitative data. Nees and Walker [2007] offered further discussion of this topic; they argued that the frequency mapping allows for the emergence of patterns in data. Much like the primary display advantage of visual graphs lies in their ability to efficiently communicate otherwise unnoticed patterns based on Gestalt perceptual grouping phenomena, the perceptual grouping of tones is susceptible to the formation of Gestalts based upon tonal relationships (see, e.g., Bregman [1990]; Deutsch and Feroe [1981]; M. R. Jones [1976]).

An early study of auditory graphs found them to be comparable in efficacy to the tactile displays traditionally used to present quantitative information to the blind [Mansur et al. 1985]. Tactile data representations yielded slightly more accurate responses, while auditory graphs resulted in faster reaction times. Flowers and Hauer [1993] found that auditory graphical representations were better than visual representations of the same data for imparting information regarding central tendency and range. They later demonstrated that participants categorized sets of visual graphs and their auditory counterparts as perceptually similar or dissimilar along the same dimensions, namely, shape, slope, and degree of linearity [1995]. Another study by Flowers, Buhman, and Turnage [1997] gave further evidence for the comparability of auditory and visual presentations of scatterplots. Participants estimated the same Pearson correlation r value for visual scatterplots and their corresponding auditory graphs. More recently, Bonebright et al. [2001] determined that, in general, participants were able to match an auditory graph to a visual line graph or scatterplot of the same data. Brown and Brewster [2003] found that people could produce a visual rendition of a graph that was over 80% accurate (on average) after hearing an auditory presentation.

1.3 Complexity of Graphical Displays

Despite the promise of auditory graphs as a viable information display, studies of the data properties that contribute to auditory graph complexity are lacking. The extensive literature on visual graph comprehension, however, can potentially offer insight into those factors affecting graph complexity. Several models and theories of graph comprehension have been proposed [Carpenter and Shah 1998; Carswell 1992; Gillan and Lewis 1994; Kosslyn 1989; Pinker 1990]. Such theories and models cannot necessarily be directly applied to auditory graphs, but auditory graph researchers would be remiss to ignore the extensive literature and theory on visual graph comprehension.

An early study of visual graph comprehension found that performance on a trend identification task decreased as the number of data points portrayed in a line graph increased [Schutz 1961], which suggested that the number of data points in a graph contributes to complexity. Carswell et al. [1993;

Carswell and Ramzy 1997] have operationally defined graphical complexity as a function of the graph's number of data points, symmetry, and number of trend reversals, whereby more data points, asymmetrical functions, and more trend reversals were factors associated with greater graphical complexity. Meyer [2000; Meyer et al. 1997] defined graphical complexity according to the number of data points in the graph (more data points being more complex), the number of data series concurrently presented (where increasing the number of series leads to greater complexity), and the graphed data's pattern (e.g., the presence of meaningful data trends reduces complexity). Similarly, Friel et al. [2001] claimed that graphical complexity "refers not only to the number of data items or categories, but also to the kinds of data types (e.g., discrete versus continuous), the spread and variation within the data set, and so on" (from paragraph 3 of their section "Progression of Graphs for Instruction"). Clearly, as Friel et al.'s "and so on" conclusion elucidates, even in the well-established literature of visual graph comprehension, descriptions of the factors contributing to the complexity of graphs vary, although researchers seem to agree that the number of data points, the number of data series, and the trend and shape of data patterns should impact graph comprehension.

To date, auditory graph studies have not manipulated fundamental characteristics of sonified data to investigate the data properties that contribute to complexity. Previous research, however, has suggested some data characteristics that might contribute to difficulty of comprehension. Consistent with suggestions from visual graph literature, Bonebright et al. [2001] found that auditory graphs portraying two data series (where data series were separated spatially such that one data series was presented simultaneously to each ear) were more difficult and took longer to match to their visual counterparts. They also found that more disperse scatterplots were more difficult to accurately match to their visual representations. Roth et al. [2002] found that participants had great difficulty (only a 25% rate of success) at identifying an auditory line graph as linear increasing; most participants mistook this function for a parabola. Performance for identifying auditory representations of parabola and sine wave functions was better (both had 75% rates of success). Roth et al.'s curious findings regarding poor performance for the identification of a simple linear increasing function, however, may be an artifact of their sonification technique. They mapped y-axis values not only to frequency changes, but also to changes in spatial elevation in the headphones. Although such redundant mappings are generally believed to be helpful in sonification [Kramer 1994], auditory spatial elevation judgments are known to be susceptible to inaccuracies (see, for example, Folds [2006]; Wenzel et al. [1993]). It seems likely that the unusual choice of a secondary spatial mapping led to the decrement in performance observed for simpler functions, as no theory of graph or auditory pattern comprehension would predict this pattern of results for a frequency mapping alone. Furthermore, previous auditory graph research had found that participants could distinguish linear from exponential functions with 84% accuracy [Mansur et al. 1985].

The studies mentioned above may offer insight into properties of data that might contribute to graph comprehension difficulty, but the majority of published auditory graph studies focus on comparisons across modalities or evaluations of different methods of sonification. No investigations that systematically manipulate data properties have been performed and meta-analytic comparisons of different properties of sonified data sets that contribute to complexity (such as the number of data points, trend reversals, or data series displayed) are not possible across studies.

1.4 Considerations from Auditory Pattern Perception Literature

While the literature specific to auditory graphs is sparse, stimuli that resemble auditory graphs have been examined in other contexts in psychology and related disciplines, such as information theory and auditory pattern perception. Although the study of the perception of frequency in time is not new, sonification and auditory graph studies have required listeners to perform novel *tasks* with sounds (e.g., data exploration and monitoring; see Kramer [1994]) as compared to past research on basic auditory perception. Despite the unique tasks required of auditory graph listeners, existing theoretical approaches may provide valuable insights and make relevant predictions regarding the perception of auditory graphs.

Martin [1972] suggested that rhythm—the "temporal patterning" (p. 487) of events in the environment—was central to information processing. He proposed that many serial stimuli (like speech, music, sequences of tones, etc.) are structured patterns in time that can be described in terms of a hierarchy of relative time and relative accents, which may involve nontemporal aspects of the stimulus. Martin posited a role for the cyclical focusing of attention in anticipation of rhythmic events and suggested that certain stimulus configurations—namely those that propagate rhythms that follow a listener's expectancies within the context of the hierarchy—were more amenable to processing than others.

M.R. Jones [1976] greatly expanded upon the work of Martin [1972] by elaborating a rhythmic theory that was derived from studies of auditory perception. The relevant dimensions of Jones' theory were frequency, loudness, and time, and the interactive nature of relational changes in the frequency and loudness dimensions of auditory stimuli as they occur in time was emphasized. Jones theorized that certain stimulus configurations would be more favorable to processing than others and the relative ease of comprehension of the patterns was described along a continuum ranging from nominal to interval perceptual relations. For sounds that are nominally related, the listener can perceive only that the sounds are the same or different. Ordinal relations allow for the perception of one sound as having higher or lower frequency than its comparator. Finally, interval relationships between tones are perceived as having both a direction and a magnitude. While, objectively, all sound relations could be described in interval terms (i.e., direction and magnitude of frequency differences between sounds can always be objectively quantified), rhythmic theory described the subjective, perceptual experience of listeners as a function of the interactions of dimensions, of which frequency and time are most relevant to the current discussion. Jones' theory suggested a continuum leading up to an upper limit for veridical perception of frequency changes in time (also see Bregman [1990]), whereby greater and more irregular frequency changes per unit of time in an auditory pattern are more difficult to process than unidirectional, regular frequency changes per unit of time.

Deutsch and Feroe [1981] proposed a model of the perception of tone sequences that was very similar in concept to M.R. Jones' [1976] theory. Deutsch and Feroe suggested that the initial perceptual grouping of sequences of tones is determined by Gestalt principles like proximity (in either frequency or time) and good continuation (such as when a series of sounds consistently increase in frequency). Their model predicted that patterns that better follow Gestalt grouping principles are easier to perceive. Deutsch and Feroe also suggested that large and more frequent changes in frequency intervals from tone to tone are detrimental to grouping. Tones that match anticipated patterns strengthen the representation of the sequence, and mismatches serve to elaborate the representation, often at higher levels in the hierarchy (provided a hierarchical rhythmic structure is present).

Auditory pattern perception theories (i.e., Deutsch and Feroe [1981]; M. R. Jones [1976]) make explicit predictions about the ease with which auditory patterns can be perceived. Specifically, they suggest that easily perceivable patterns are characterized by frequency changes over time that allow for the grouping of tones based on Gestalt principals, such as temporal proximity, frequency proximity, and expectations of good continuation of the pattern. Simple (e.g., monotonically increasing) data patterns in auditory graphs, then, should allow for better performance of auditory graphing tasks than more complex data patterns with more trend reversals per unit of time. Despite the fairly straightforward predictions of these theories regarding ease and difficulty in auditory pattern perception, these theories did not conceive of auditory patterns as graphs, nor were they intended to address the exact types of *tasks* (e.g., point estimation and trend identification) that are required of auditory graph listeners. The auditory perception literature suggests, however, that greater and more irregular frequency changes per unit of time should generally result in sound patterns that are more difficult to perceive, and these predictions should translate to performance tasks with auditory graphs. The systematic examination of the fundamental data properties that affect auditory graph comprehension should provide valuable insights for sonification researchers. Kosslyn [1989] suggested that visual graphs provide for the automatic perception of some data features, as Gestalt patterns (based on spatial proximity, good continuation, etc.) can emerge from visual graphs of data. Certain data features (e.g., a simple increasing trend) likewise may be readily perceivable from an auditory graph, while more complex data features may make information extraction more difficult. The identification of those properties of the data that may affect auditory graph comprehension should permit auditory graph designers and listeners to better understand and predict the attributes of data (e.g., patterns) that will be easily perceived from an auditory graph.

1.5 The Current Research

In the present pair of studies, we manipulated two fundamental properties of auditory graph stimuli—the density of data points and number of trend reversals displayed—to examine the effects of graph complexity for both point-estimation (Experiment 1) and trend-identification (Experiment 2) tasks. For both studies, participants listened to auditory graph stimuli that depicted the price of a fictitious stock over the course of a 10-hr trading day. Participants in the point-estimation experiment were asked to estimate the price of the stock at a given hour of the day on each trial. For the trend-identification experiment, participants identified the stock price's direction of change (i.e., increasing or decreasing) for a given 1-hr segment of the trading day.

Of note, these studies examined auditory line graphs, where no more than one y-axis value was displayed for a given x-axis value. Although auditory graph researchers have investigated auditory scatterplots [Bonebright et al. 2001; Flowers et al. 1997], box-whisker plots [Flowers and Hauer 1992; Peres and Lane 2003, 2005], histograms [Flowers and Hauer 1993], and tabular data [Stockman et al. 2005], the majority of auditory graph research has examined auditory line graphs [Bonebright et al. 2001; Brewster and Murray 2000; L. M. Brown et al. 2002; L. M. Brown and Brewster 2003; Flowers and Hauer 1995; Mansur et al. 1985; Roth et al. 2002; Smith and Walker 2002, 2005; Turnage et al. 1996; Walker and Nees 2005b]. This is not surprising, considering that line graphs accounted for 72.5% of graphs appearing in academic journals and 50.1% of all graphs appearing in newspapers in the sample obtained by Zacks et al. [2002]. Likewise, Peden and Hausmann [2000] found that 63% of the graphs appearing in introductory psychology textbooks were line graphs.

We frame our hypotheses with respect to the tasks studied: point estimation (Experiment 1) and trend identification (Experiment 2). Although many different tasks can be required of graph users, the identification of trends and the estimation of a y-axis value for a given x-axis value are common activities that are representative of some of the typical graph user's basic information needs. Other data attributes such as symmetry and the initial direction of change were held constant or counterbalanced in these studies.

1.6 Hypotheses

1.6.1 *Hypothesis 1a: Point Estimation and Data Density.* For the point-estimation task (Experiment 1), performance was predicted to decline as data density increased (i.e., as more data points were presented). This prediction was consistent with empirical findings and theoretical predictions from visual graph comprehension literature (e.g., Carswell et al. [1993]; Carswell and Ramzy [1997]; Friel et al. [2001]; Meyer [2000]; Meyer et al. [1997]; Schutz [1961]) and makes intuitive sense, because with more data points the listener must select the target from a larger array of discrete tones. This prediction is somewhat in conflict with the predictions of rhythmic theories like those of Martin [1972], which predicted no perceptual decrements no matter the density of tones per second in auditory patterns, and of M.R. Jones, [1976], which predicted a decrement under circumstances of fast rates of presentation coupled with extreme frequency jumps (rather than as a function of the rate of presentation per se). Neither rhythmic theory, however, was intended to predict performance for this particular task (point estimation) with auditory patterns.

1.6.2 *Hypothesis 1b: Point Estimation and Trend Reversals.* Performance on the point-estimation task was also predicted to decline as the number of trend reversals in the displayed data increased. This prediction was consistent with empirical findings and theoretical predictions from visual graph comprehension literature (e.g., Carswell et al. [1993]; Carswell and Ramzy [1997]). This hypothesis was also consistent with predictions regarding auditory pattern complexity proposed by the theories of M.R. Jones [1976] and Deutsch and Feroe [1981].

1.6.3 *Hypothesis 2a: Trend Identification and Trend Reversals.* Regarding trend identification with auditory graphs (Experiment 2), performance on trend identification was predicted to decrease as the number of trend reversals increased. This prediction was consistent with research in the visual graph comprehension literature that has posited a role for trend reversals in the complexity of graphical displays (e.g., Carswell et al. [1993]; Carswell and Ramzy [1997]). This prediction was also consistent with predictions regarding auditory pattern complexity proposed by the theories of M.R. Jones [1976] and Deutsch and Feroe [1981], which both suggested that directional changes in frequency create more elaborate, complicated auditory patterns than stimuli with unidirectional frequency change.

1.6.4 *Hypothesis 2b: Trend Identification Interaction of Density and Trend Reversals.* For the trend-identification task, an interaction was also predicted such that lower data density (i.e., fewer data points presented per second in the display) would result in worse trend-identification performance as the number of trend reversals increased. The lower data density auditory graphs offer less information for participants to discern the data trends in the face of increasing trend reversals; M.R. Jones [1976] generally suggested that the perception of sequences of tones is more difficult when frequency changes are greater or occur more often per unit of time.

2. EXPERIMENT 1: POINT ESTIMATION

Experiment 1 examined the effects of data density and the number of trend reversals within the sonified data on performance of a point estimation task with auditory graphs.

2.1 Method

2.1.1 *Participants.* Participants ($N = 32$; 16 males and 16 females) were recruited from undergraduate psychology courses at the Georgia Institute of Technology. Participants' mean age was 19.1 ($SD =$ 1.6) years, and the ages sampled ranged from 18 to 25 years old. Participants reported having played a musical instrument for an average of 3.5 ($SD = 3.9$) years, with 9 participants having never played an instrument and 11 having played an instrument for 5 years or more. They reported a mean of 2.9 ($SD =$ 3.5) years of formal musical training (i.e., private or class instruction), and a mean of 3.19 ($SD = 3.96$) years of experience with reading musical notation. Participants had a mean of 0.3 ($SD = 0.7$) years of experience with stock trading or closely following the stock market. They also reported having taken a mean of 2.9 ($SD = 2.5$) college or advanced placement-level business or economics courses. The two latter demographic questions were included as indicators of prior domain knowledge for the sonified data, which was the price of a fictional stock.

2.1.2 *Apparatus.* Visual presentations (such as instructions and text presentations of questions during trials) were made on a 17-in. (43.2-cm) LCD computer monitor. Auditory presentations were delivered via Sennheiser HD 202 headphones, which were adjusted by the participant to a comfortable fit. Listening volume was approximately 65 dB SPL. All presentations of stimuli and data collection were accomplished with the Macromedia Director MX 2004 software package.

2.1.3 *Data Sets for Stimuli.* Auditory graphs depicted the price of a stock in dollars as it varied (within a range of 6 to 106 dollars) over the course of a 10-hr trading day that opened at 8 a.m. and closed at 6 p.m. Fictional stock price data have been used in past research (e.g., Smith and Walker [2005]; Walker and Nees [2005b]), because they represent a relatively generic domain that should be accessible to naïve subjects with no specialized expertise in the area. To further ensure that domain knowledge did not influence performance for tasks in the current study, participants were given a brief overview of the task with information regarding the domain (i.e., stocks have monetary values that fluctuate during a day of market trading). No additional prior knowledge of the domain was necessary for the graphing tasks in the current study.

Data density was manipulated at four levels that offered a psychologically and practically relevant range of stimuli: one data point per second, two data points per second, four data points per second, and eight data points per second. Each successive level, therefore, doubled the tempo of the previous level (and, likewise, roughly doubled the perceived tempo of the previous level; see Walker [2002; 2007]). With regard to the range of stimuli employed here, stimulus interonset intervals (IOIs) under 1800 ms promote perceptual grouping; with longer intervals items may be perceived independently rather than as members of a sequence (see Fraisse [1978; 1982]). Perception of tones as a coherent auditory graph, therefore, may fail if data are not presented at a minimally sufficient rate and the lowest data density employed here falls well within the limit of grouping by IOIs for tones.

Furthermore, from a practical perspective, auditory graphs will need to be designed such that a listener can explore the data in a reasonable amount of time. For example, rhythmic sequences, characterized by presentation rates around one item per second, have been characterized as perceptually "slow," while about five to six items per second are perceived as "fast" [Palomaki 2006] and, in studies of tempo, researchers have operationally defined stimuli with at or near four items per second as "fast" [M. R. Jones et al. 2006] and at or near 10 items per second as "very fast" [Drake and Botte 1993]. Of further interest regarding reasonable upper limits for the temporal presentation of data in auditory graphs, the threshold for determining the order of temporally presented stimuli (i.e., being able to perceive which item preceded adjacent items in a series) has been shown to range from 20 to 100 ms [Fraisse 1978] and an auditory graph listener would be served poorly by a graph where this "threshold of succession" was ambiguous. The current study's fastest rate of presentation, while falling within a presentation rate that is perceived as "fast," was well below the rate whereby succession of items was indeterminate in previous research. Another important consideration involved the duration of discrete, individual tones, which necessarily decreased as more tones were added per unit of time. Early research [Turnbull 1944] suggested that frequency discrimination deteriorated rapidly as the length of a tone fell below durations of around 100 ms. The current study used tones of a duration well above the threshold for failures of frequency discrimination because of tonal duration.

The second experimental variable involved the number of trend reversals presented in the auditory graph stimuli. The four levels of the trend-reversals variable were operationally defined as zero, one, two, and three trend reversals in the data. Auditory graphs with zero-trend reversals represented data that either increased or decreased monotonically across the entire trading day. Graphs with one trend reversal rose for the first half of the trading day, then fell, or vice versa. Graphs with two and three trend reversals assumed two and three changes in the direction of the stock price data trend, respectively.

Figure 1 offers visual depictions of each of the combinations of data density and trend reversals. Of note, for each combination of the two independent variables, auditory graphs could be constructed such

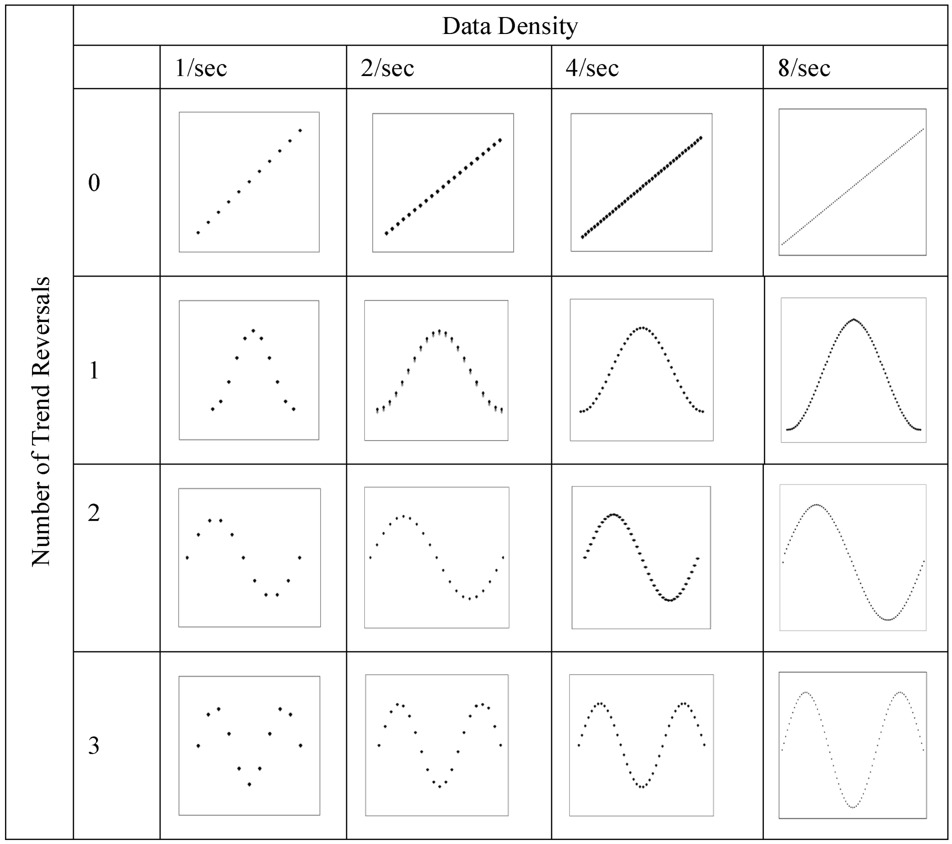

Fig. 1. Visual depictions of the 16 factorial combinations of stimuli with initially increasing data values. Note that these are schematic depictions only; see Section 2.1.4 for specific details on the design of auditory graph stimuli.

that the data in the graphs initially increase (such as in Figure 1) or initially decrease. For a given factorial combination of data density and trend reversals from the 16 different stimulus combinations, participants experienced only graphs that initially increased or only graphs that initially decreased within a stimulus combination cell. Across stimulus conditions, a given participant experienced 50% of trials with initially increasing graphs and 50% of trials with initially decreasing graphs. For example, for data density of one data point per second coupled with zero trend reversals, half of the participants experienced linear increasing graphs; the other half experienced linear decreasing graphs.

2.1.4 *Auditory Graph Stimuli Design.* The price of the stock in dollars (on the y-axis) was represented by discrete tones that changed in frequency as the price changed, while each hour of the trading day (on the visual x-axis) was mapped to 1 s in time. The duration of individual tones was held constant at 125 ms per tone, with 10 ms onset and offset ramps. All auditory graphs were 10.125 s in duration. Data for auditory graphs were sonified as MIDI data files using the *Sonification Sandbox* program (see Walker and Cothran [2003]; Walker and Lowey [2004]), converted to .WAV files, and exported to Audacity version 1.3.0b for track mixing. All sounds presented the same signal to both left and right headphone channels (i.e., a center mix without stereo panning was employed). Within the *Sonification Sandbox* program, stock data were represented with the MIDI instrument bank's piano timbre. Scaling anchors were assigned for maximum and minimum values in a data set. The minimum data value ($6) was assigned MIDI note G2, whose frequency was 98 Hz, and the maximum data value ($106) corresponded to MIDI note B6, whose frequency was 1979.5 Hz. The full range of the chromatic scale from note G2 to note B6 was used in each auditory graph and each graph attained maximum and minimum data values. A positive polarity mapping was employed and data values in between the maximum and minimum were assigned MIDI notes on an exact scale (i.e., in the event that data values fell between notes on the chromatic scale, the tones were adjusted to represent the exact frequency of the data point on the scale).

2.1.5 *Auditory Context for X and Y Axes.* Auditory context in the form of y-axis reference tones and x-axis click tracks has been shown to generally provide cues that aid in the performance of point estimation tasks with auditory graphs [Smith and Walker 2002; 2005]. Despite the potential for auditory context to aid auditory graph comprehension, both y-axis reference tones and x-axis click tracks were omitted in this study, as concurrent auditory context represented a potential confound when implemented with the manipulations of the current studies. Regarding y-axis reference tones, the most useful reference tones alternate (from maximum to minimum references) as the data trends change [Smith and Walker 2002, 2005]. Given that the current study investigated the impact of trend reversals, the number of reference tone alternations across stimuli would change as a function of trend reversals, perhaps producing differential effects with regard to auditory stream segregation [Bregman 1990] and contextual benefit.

Furthermore, considerable research has indicated that rhythmic perception is hierarchical, with lowest nodes at the level of individual discrete sounds, which are grouped according to accents (i.e., rhythmic beats), etc., proceeding upward to higher-order, more complex temporal organizations (see, for example, Deutsch and Feroe [1981], Fraisse [1978, 1982]; M. R. Jones [1976]; Povel and Essens [1985]). Correspondingly, auditory graph research has suggested that x-axis rhythmic context (in the form of clicks or beats) was more effective when the accents were placed at a density that was less than the density of the actual discrete data points [Smith and Walker 2002, 2005]. In other words, a click track helped little with temporal organization when the clicks coincided with each discrete data point (and the clicks offered only redundant information), but the click track was beneficial when it played, for example, once every two data points.

The manipulations of the current studies were such that both y-axis and x-axis auditory context could not be held constant across the manipulations, as the contextual manipulations would potentially offer more benefit for some conditions than others. Instead, context was provided via the instructions and preexperimental practice. Participants were familiarized with the scaling of data (i.e., which frequencies represented the highest and lowest prices of the day) as well as the mapping of stock price to the increasing and decreasing frequencies of tones. For all trials, participants were told the opening price of the stock (e.g., $50) as an additional contextual cue. In other words, although concurrent auditory context was not used, participants were given other contextual cues (that *could* be held constant across all conditions) to help them perform the graphing tasks.

2.1.6 *Procedure and Task.* Informed consent was obtained from participants before any procedures were performed. Because of the novelty of the display, all participants experienced a brief (approximately 10–15 min), self-paced, conceptual background presentation (see Smith and Walker [2005]; Walker and Nees [2005b]) that gave an overview of auditory graphs (e.g., "What is an auditory graph?"), as well as instructions for the point-estimation task. The background presentation included information about the different data densities employed in the study, and they were instructed that 1 s in the auditory graph represented 1 hr of the trading day, regardless of the density. Participants were also given part-task practice on important component steps of the point-estimation task during the introductory presentation, followed by a full set (16 trials) of whole-task practice with feedback on the experimental point-estimation task. During these practice trials, participants heard auditory graph stimuli that were equivalent to the stimuli presented during test trials (with respect to the data density and the number of trend reversals) and the practice trials sampled the entire range of stimulus combinations that were used in the test trials. At no time during practice, however, did a participant hear any stimulus that was later used in a test trial. This was possible because of the counterbalancing of graphs with increasing and decreasing initial deflections across stimulus combinations. A person who would later experience linear increasing *experimental* trials (for a given combination of data density and trend reversals) was given linear decreasing examples (of the same given combination of density and trend reversals) during *practice* trials. Both part- and whole-task training have been shown to improve performance on the point-estimation task [Smith and Walker 2005; Walker and Nees 2005b], so the combined effects of both part- and whole-task practice before testing helped to reduce effects of unfamiliarity with the display that might otherwise have been present in early test trials. Participants had the opportunity to ask questions to resolve any questions about the nature of the auditory graph stimuli and the task.

The 4 (density) x 4 (trend reversals) design resulted in 16 stimulus combinations. Participants therefore experienced 11 sets of 16 experimental trials, with each set consisting of randomly interleaved trials, one from each of the 16 data density and trend-reversal stimulus combinations. Over the course of all 11 sets, participants were asked to identify the price of the stock for each hour of the trading day (8 A.M.–6 P.M.) for each stimulus condition (16 combinations of data densities and trend reversals) in a random order.

Individual trials began with a visual text presentation of the test question (e.g., "What is the price of the stock at 10 A.M.?") followed by the presentation of the auditory graph. To provide a baseline reference, participants were told the opening price of the stock for each auditory graph. A task analysis (see Smith and Walker [2005]) has shown that the point-estimation task requires the listener to: (1) listen to the auditory graph; (2) determine which part of the auditory graph corresponds to the queried hour of the day (i.e., a temporal interval division task); (3) perform a magnitude estimation task with respect to the perceived pitch of the tone; and (4) compare the pitch at the queried time to the known pitch of the opening price and assign a quantitative value. Participants were permitted to listen to the auditory graph as many times as needed before responding; the next trial began after a response was recorded with the computer keyboard. Participants were given mandatory 5-min breaks between sets 4 and 5 and between sets 8 and 9. In an effort to promote engagement with the task, participants were given feedback about their performance for the set (as a whole) at the end of each of the 11 sets of experimental trials (e.g., "You were within 5 dollars (or less) of the correct stock price on 8 out of 16 trials for this set."), but no specific feedback was given after *individual* experimental trials.

2.1.7 *Dependent Variables.* The primary dependent variable of Experiment 1 was operationally defined as the root mean squared (RMS) error in dollars of participants' responses to the point estimation trials for each of the 16 stimulus combinations. For a more detailed analysis of the use of RMS error in point estimation sonification tasks, see Smith and Walker [2005]. The mean number of times participants listened to graphs for each combination was a second dependent variable of interest.

2.2 Results

One outlier datum was removed from the analyses. One participant (presumably accidentally) responded with an estimated stock price of 878 dollars for the 10 A.M. question for the three trend reversals, eight data points per second stimulus combination. Participants were explicitly told the data ranged up to only 106 dollars and the participant gave no extreme responses for any other questions. This datum was not included in analyses and the participant's RMS error for that stimulus combination was computed as a mean out of 10 trials instead of 11.

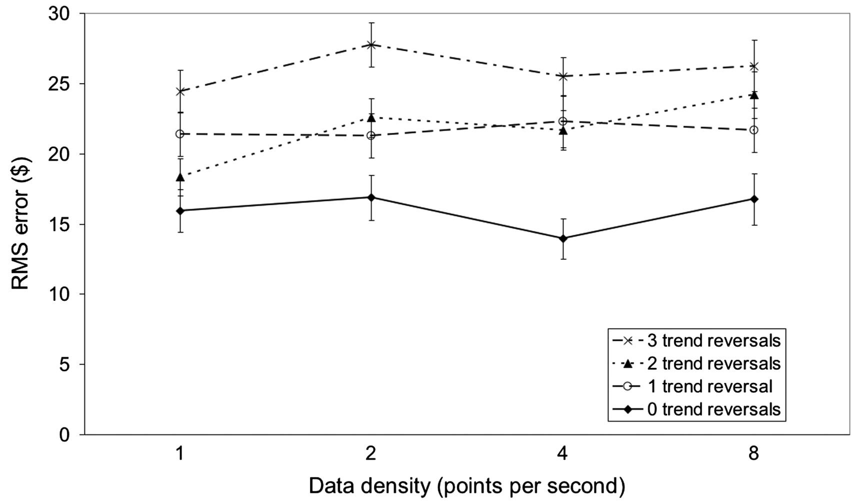

Fig. 2. Overall results of the point-estimation task (Experiment 1) for the RMS error-dependent variable. Note that higher error indicates worse performance; error bars represent standard error.

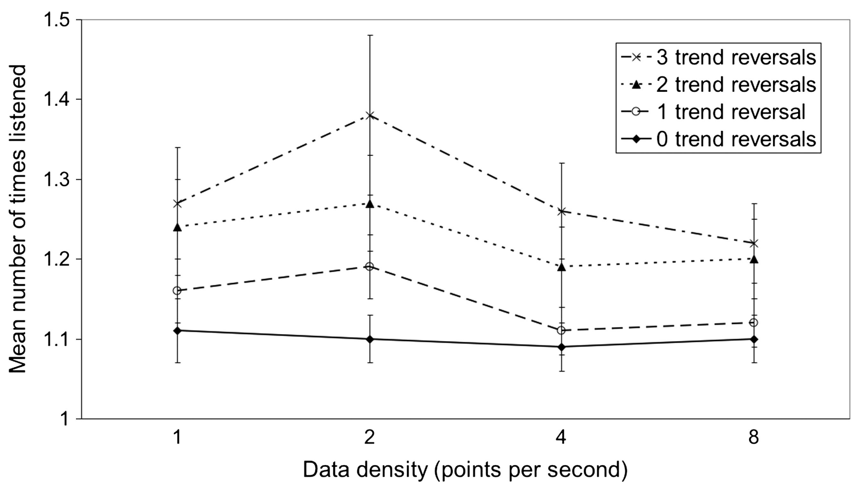

Fig. 3. Overall results of the point-estimation task (Experiment 1) for the mean number of times listened dependent variable. Error bars represent standard error.

To check the assumption that the initial deflection of the graphed data (either increasing or decreasing) had no impact on performance, a series of one-way analyses of variance (ANOVAs) was performed (for each combination of data densities and trend reversals) with the initial deflection of the graph as the independent variable. The Bonferroni procedure was used to protect family-wise alpha across this set of analyses. As was assumed, results for both dependent variables (RMS error and the mean number of times a graph was played) showed no significant difference with regard to the initial deflection of the graph for any stimulus combination; for further analyses the data were collapsed across the initial direction of the graph. Overall results for the point-estimation study are depicted in Figures 2 (RMS error) and 3 (mean number of times listened).

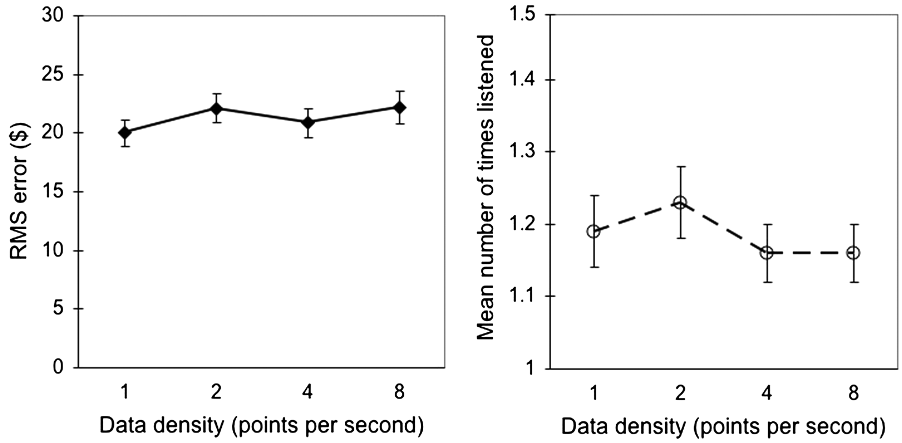

Fig. 4. Results of the point-estimation task (Experiment 1) for the significant main effect of data density on both dependent variables. Error bars represent standard error.

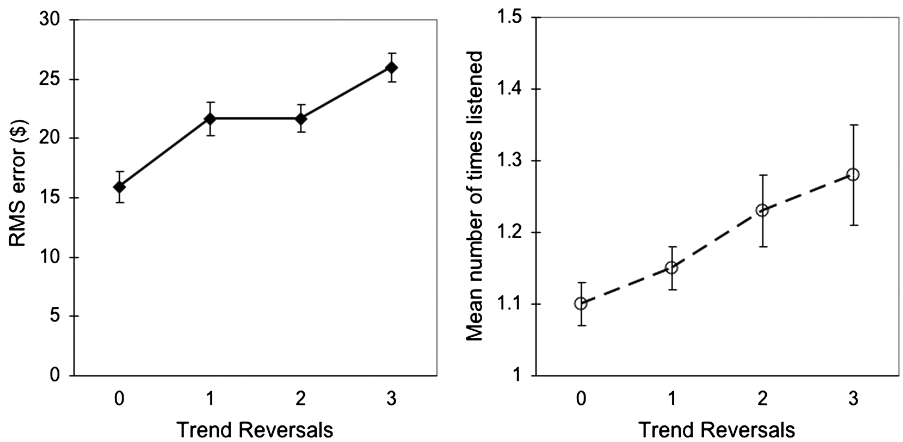

Fig. 5. Results of the point-estimation task (Experiment 1) for the significant main effect of trend reversals on both dependent variables. Error bars represent standard error.

Data were analyzed with 4 (data density) x 4 (trend reversals) repeated measures analyses of variance (ANOVAs); Huynh–Feldt corrections were employed in all analyses where sphericity assumptions were violated. The main effect of density was significant on both RMS error scores [F (3, 93) = 3.08, $p = .03$, partial $\eta^2 = .09$] and the mean number of times listened to each graph [F (2.0, 62.2) = 6.86, $p = .002$, partial $\eta^2 = .18$]. Results for the main effect of data density (collapsed across trend reversals) are depicted in Figure 4.

Results for the main effect of trend reversals (collapsed across data density) are depicted in Figure 5. For the trend reversals independent variable, the ANOVAs showed a significant main effect on both RMS error scores [$F(3, 93) = 41.09$, $p < .001$, partial $\eta^2 = .57$] and the mean number of times listened to each graph [$F(1.6, 49.6) = 10.27$, $p < .001$, partial $\eta^2 = .25$]. The interaction of data density and trend reversals was not significant for either RMS error [$F(9, 279) = 1.87$, $p = .06$] or the mean number of times listened [$F(4.9, 151.6) = 1.60$, $p = .17$] in Experiment 1.

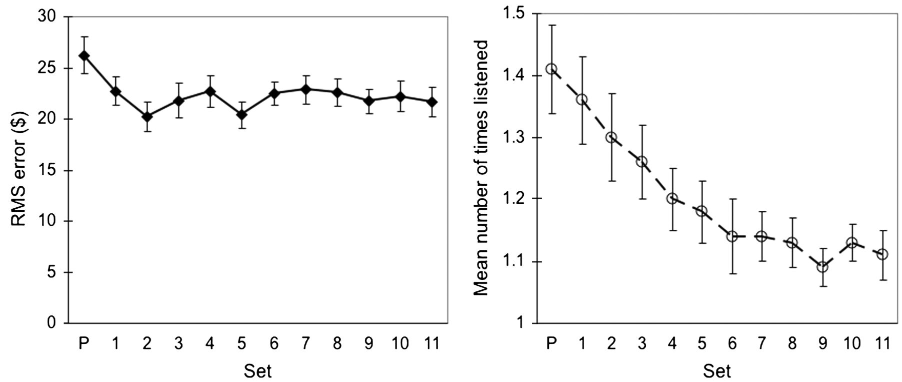

Fig. 6. Results over time for the point-estimation task (Experiment 1). The first set featured practice trials with feedback, while the rest of the sets were experimental trials with no specific feedback. Error bars represent standard error.

An exploratory analysis was conducted to determine if performance changed across presentations of sets in time (collapsed across data density and trend reversals). As described above, each presentation set consisted of 16 randomly interleaved trials, each with one presentation of each factorial stimulus combination. To investigate practice effects, another repeated measures ANOVA was employed with the RMS error for each set and the mean number of times listened to the graphs in the set as the dependent variables. Results across sets (i.e., time) in Experiment 1 are depicted in Figure 6. This analysis indicated effects of set on both RMS error [$F(11, 341) = 2.65$, $p = .003$, partial $\eta^2 = .079$] and the mean number of times listened to each graph [$F(4.3, 133.1) = 11.58$, $p < .001$, partial $\eta^2 = .27$].

Finally, demographic variables related to musical experience and experience with stock trading, math, and business or economics courses were correlated with each other, as well as with a gross measure of overall performance—the average RMS error across the entire study. Spearman's rho was employed as a result of the nonnormal distribution of the demographic variables. None of the demographic questions were significantly associated with performance on the point-estimation task.

2.3 Discussion

Hypothesis 1a, regarding data density, predicted an increase in RMS error as the number of data points per second increased. The data density manipulation showed a small, but significant, effect on both RMS error scores and the mean number of times listened to graphs. As depicted in Figure 4, mean differences in RMS error were small across manipulations of data density and this pattern of results does not conclusively confirm or disconfirm Hypothesis 1a.

Results for the main effect of trend reversals for the point-estimation task generally confirmed Hypothesis 1b that performance would decrease as the number of trend reversals increased. As depicted in Figure 5, the best performance (i.e., lowest RMS error and lowest mean number of times listened for the graphs) occurred when there were no trend reversals; the worst performance occurred with three trend reversals in the data.

Together these results suggest that the number of trend reversals played a large role in participants' ability to perform the point-estimation task, whereas the number of data points presented per second had a less substantial influence on performance outcomes. The results regarding data density may be of less practical than statistical significance, as the mean differences across all conditions were quite small and account for relatively little performance variance as compared to the trend-reversals manipulation. Further research is needed to clarify the extent to which data density can impact performance with point-estimation tasks in auditory graphs. Performance (as measured by both RMS error and the mean number of times listened to each graph) clearly decreased as the number of trend reversals increased from 0 to 3. The data for RMS error showed a plateau for performance with 1 and 2 trend reversals (see Figure 5). Although this plateau is not evident in the data for the mean number of times listened, it warrants further investigation in future studies.

The exploratory analyses to look for effects of set over time on the dependent variables suggested practice effects for both dependent variables. While the mean difference for RMS error over time are small (see Figure 6), the plot of the number of times listened suggested that participants listened to the auditory graph stimuli fewer times, on average, at the end of the study. Although the current data cannot conclusively explain this effect, the combination of significantly fewer mean times listened in set 11 without a corresponding drop in accuracy (RMS error) in later trials possibly suggests that by the end of the study participants did not need to hear the auditory graph stimuli as many times to maintain task performance.

3. EXPERIMENT 2: TREND IDENTIFICATION

Experiment 2 examined the effects of data density and the number of data trend reversals on performance of a *trend-identification task* with auditory graphs. A new sample of participants was recruited, but the apparatus, stimuli, and experimental manipulations were the same as those in Experiment 1. The only substantial difference in methodology between Experiments 1 and 2 was the task; participants in Experiment 2 were asked to identify local trends for each stimulus combination, as described below.

3.1 Method

3.1.1 *Participants.* Participants ($N = 32$; 22 males and 10 females) were recruited from undergraduate psychology courses at the Georgia Institute of Technology, and none had participated in Experiment 1. Participants' mean age was 19.0 ($SD = 1.4$) years, and the ages sampled ranged from 18 to 24 years old. Participants reported having played a musical instrument for an average of 4.8 ($SD = 3.5$) years, with 3 participants having never played an instrument and 16 having played an instrument for 5 years or more. Participants reported a mean of 3.9 ($SD = 3.0$) years of formal musical training (i.e., private or class instruction) and a mean of 4.9 ($SD = 3.8$) years of experience with reading musical notation. Participants had a mean of 0.5 ($SD = 1.1$) years of experience with stock trading or closely following the stock market. They also reported having taken a mean of 2.6 ($SD = 1.8$) college or advanced-placement-level math courses and an average of 1.0 ($SD = 1.0$) college or advanced-placement-level business or economics courses.

3.1.2 *Procedure and Task: Differences between Experiments 1 and 2.* The procedure and task for Experiment 2 were the same as Experiment 1 with a few notable exceptions. The introductory instructional materials and practice trials in Experiment 2 were like those in Experiment 1, except the session was tailored to help participants to perform trend identification rather than point estimation. In Experiment 2, participants were asked to identify the local *trend* of the stock between all 10 successive hours of the trading day (e.g., "Did the price of the stock increase, decrease, or stay the same between 10 A.M. and 11 A.M.?") for each stimulus condition in a random order. Participants were limited to three response choices for trend: increased, decreased, or stayed the same.

3.1.3 *Dependent Variables.* The dependent variable of Experiment 2 was operationally defined as the percentage of correct responses on the local trend identification task for each of the 16 stimulus combinations. As before, the mean number of times participants listened to graphs for each set was a second dependent variable of interest.

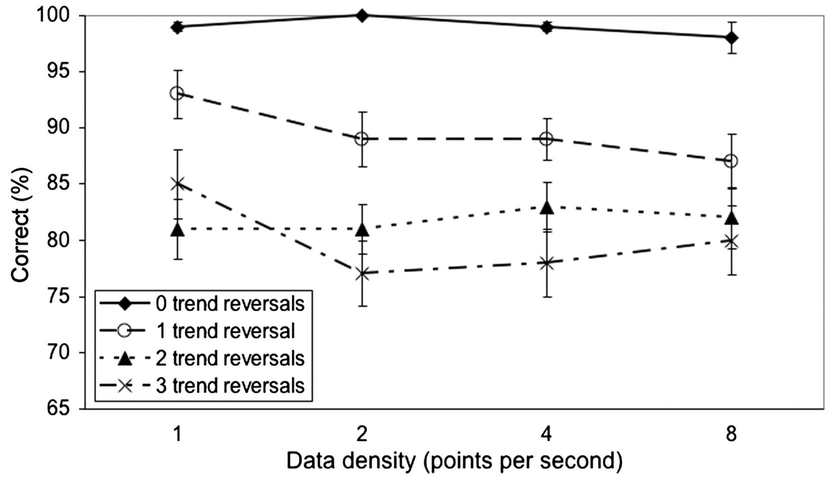

Fig. 7. Overall results of the trend-identification task (Experiment 2) for the percentage correct variable. Error bars represent standard error.

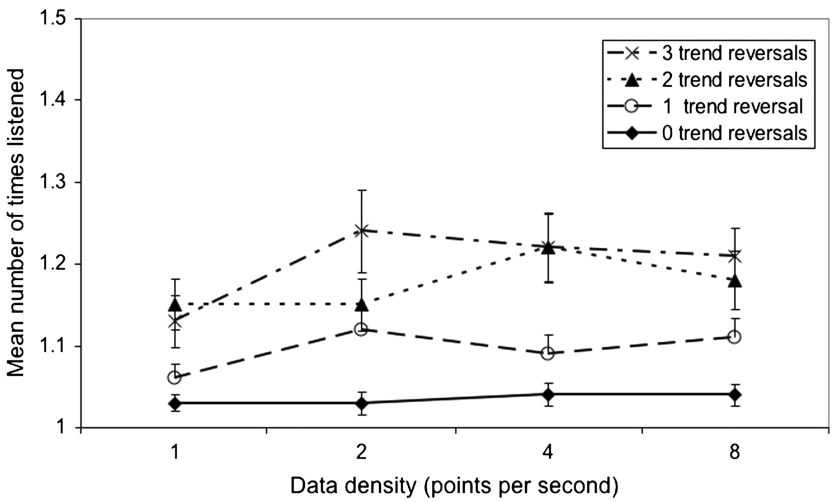

Fig. 8. Overall results of the trend-identification task (Experiment 2) for the mean number of times listened dependent variable. Error bars represent standard error.

3.2 Results

For Experiment 2, the first analyses again checked the assumption that the initial deflection of the graphed data (either increasing or decreasing) had no impact on performance with a series of one-way ANOVAs for each combination of data densities and trend reversals. As was assumed, results for both dependent variables (percentage correct and the mean number of times a graph was played) showed no significant differences with regard to the initial deflection of the graph for any stimulus combination; for further analyses the data were collapsed across this variable. Overall results for the trend-identification study are depicted in Figures 7 (for percentage correct) and 8 (for mean number of times listened).

For the primary analyses, 4 (data density) x 4 (trend reversals) repeated measures ANOVAs for each dependent variable showed a significant main effect of data density on the mean number of times

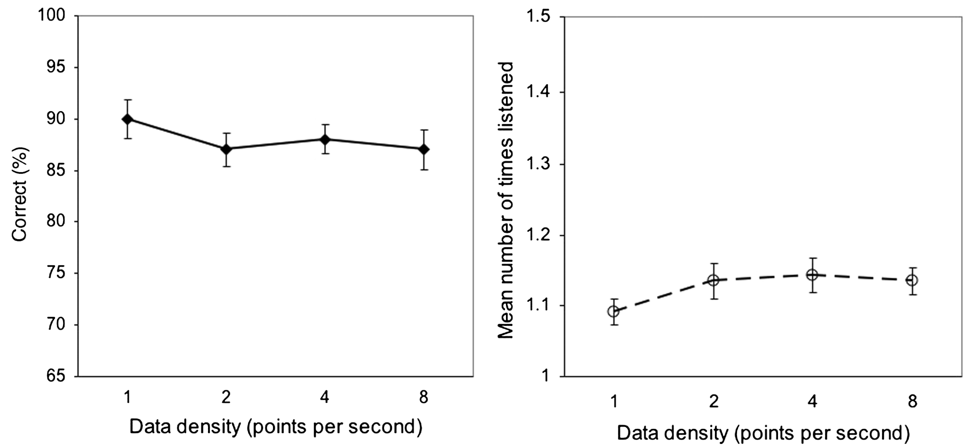

Fig. 9. Results of the trend-identification task (Experiment 2) for the main effect of data density. Error bars represent standard error.

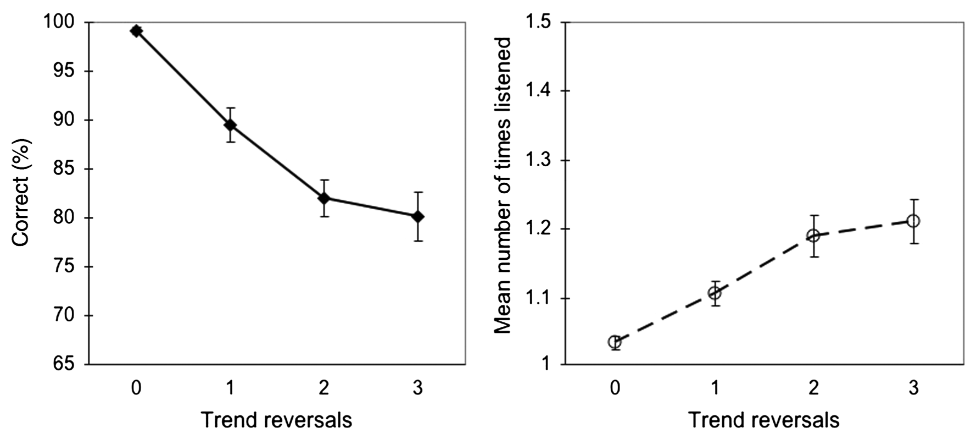

Fig. 10. Results of the trend identification task (Experiment 2) for the significant main effect of trend reversals on both dependent variables. Error bars represent standard error.

participants listened to each type of graph [$F(2.5, 77.8) = 3.80$, $p = .02$, partial $\eta^2 = .11$], but not on the percent correct[1] [$F(3, 93) = 2.04$, $p = .11$]. Results for the main effect of data density (collapsed across trend reversals) are depicted in Figure 9.

The main effect of trend reversals was significant on both percentage correct scores [$F(1.9, 58.8) = 60.78$, $p < .001$, partial $\eta^2 = .66$] and the mean number of times listened to each graph [$F(2.1, 66.0) = 23.57$, $p < .001$, partial $\eta^2 = .43$]. Results for the main effect of trend reversals (collapsed across data density) on both dependent measures are depicted in Figure 10.

[1]Proportions can be problematic in ANOVA, especially with large percentage scores such as those observed in the current study. An alternative analysis strategy involves the use of the arcsine transformation of proportions (for a discussion, see Keppel and Wickens [2004]). We analyzed the more interpretable percentage scores here. However, note that an ANOVA performed on arcsine transformed percentage scores yielded a pattern of results that was identical to those reported for our analysis of untransformed percentage scores.

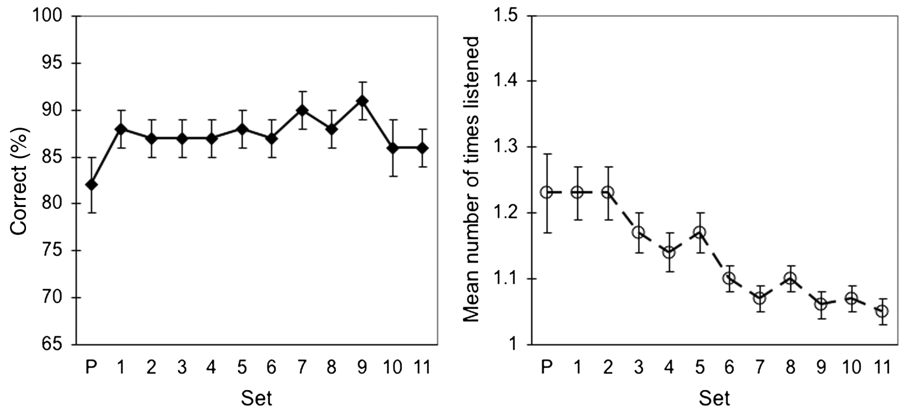

Fig. 11. Results over time for the trend-identification task (Experiment 2). The first set featured practice trials with feedback, while the rest of the sets were experimental trials with no specific feedback. Error bars represent standard error.

The interaction of data density with trend reversals was significant for percentage correct [$F(7.1, 218.7) = 2.11$, $p = .04$, partial $\eta^2 = .06$], but not for the mean number of times listened to each graph stimulus [$F(6.5, 200.2) = 1.48$, $p = .18$]. The interactions are observed in the upper and lower left corners of Figure 7.

Like in Experiment 1, a final, exploratory set of analyses was conducted to determine if performance changed across presentations of sets in time (collapsed across data density and trend reversals). Results across sets (i.e., time) in Experiment 2 are depicted in Figure 11. The ANOVAs indicated effects of set on the mean number of times listened [$F(2.7, 83.0) = 8.15$, $p < 0.001$, partial $\eta^2 = 0.21$], but not on the percentage correct [$F(10.1, 313.2) = 1.78$, $p = 0.06$].

Demographic variables related to musical experience and experience with stock trading, math, and business or economics courses were correlated with each other as well as with a gross measure of overall performance—the average percentage correct across the entire study. Spearman's rho was employed due to the nonnormal distribution of the demographic variables and, again, none of the demographic questions were significantly associated with performance on the trend-identification task.

3.3 Discussion

Figure 9 depicts the significant main effect of data density on the mean number of times listened to each graph. The figure suggests that participants needed slightly fewer mean times listened for graphs that had one data point per second as compared to other density conditions. The main effect of data density on the number of times listened was small, however, and no main effect of data density was found for the percentage correct measure. Again, more research is needed to conclusively determine the impact of data density on auditory graph comprehension, but the current results suggest that density had a small, but significant, impact such that fewer data points per second lead to slightly better performance.

A main effect of the number of trend reversals was found for both dependent measures; these are depicted in Figure 10. Performance on the task decreased concurrently for both measures, as reflected in simultaneous decreases in percentage correct and increases in the mean number of times listened to the graphs as the number of trend reversals increased.

The finding of a main effect of trend reversals for the percentage correct variable, however, should be interpreted in light of the significant interaction of data density and trend reversals for the percentage correct. (No such interaction was present for the mean number of times listened). The interaction is depicted in Figure 7. While performance with zero trend reversals remained near ceiling, moving from one to two data points per second (and across all densities, for that matter), the interaction occurred such that the introduction of one trend reversal showed a decline in performance at two data points per second. Interestingly, Figure 7 suggests that, comparing two to three trend-reversal performance from one to two points per second, performance for graphs with three trend reversals was better (than performance with only two trend reversals) at one point per second. Performance with three trend reversal graphs then dropped relatively sharply at two data points per second. This was contrary to Hypothesis 2b, which predicted an interaction characterized by performance with three trend reversals and one data point per second as the worst stimulus combination for the trend identification task.

The interaction of data density with trend reversals, coupled with the increase in the number of times participants tended to listen to a graph as data density increased from one to two data points per second, suggest that the change from one to two data points per second was problematic, particularly for conditions where graphs featured one or three trend reversals. For zero trend reversals, performance remained at ceiling across manipulations of data density, while two trend reversals in the data showed little change as density increased from one to two data points per second. The slowest rate of presentation seems to have compensated for the generally more difficult stimuli with three trend reversals, but it is unclear why the same pattern was not observed for data with two trend reversals.

Like in Experiment 1, the exploratory analysis suggested a larger effect of set on the mean number of times listened. These results for the trend-identification study generally parallel the findings from the same analyses for the point-estimation study. Participants needed fewer presentations of the auditory graph stimuli to maintain the same level of accuracy for later as compared to early trials.

4. GENERAL DISCUSSION

This study's results regarding trend reversals suggested that, as was implied in theories by both M.R. Jones [1976] and Deutsch and Feroe [1981], an auditory graph with linear increasing or decreasing data is relatively easy to comprehend. When data are represented with sound, some features of data—particularly if those features are parsimonious with regard to trend—may be easily perceived and perhaps should be characterized as auditory versions of the emergent patterns discussed by Sanderson et al. [1989] in their discussion of visual displays. This finding similarly parallels the predictions made in Kosslyn's [1989] theory of visual graph comprehension, namely, that simple patterns may be perceived automatically and efficiently in graphical representations.

The manipulation of trend reversals in the current study did not control for changes in the frequency intervals between queried data points as the number of trend reversals increased, as the frequency scaling factor was instead held constant across stimuli. Although M.R. Jones' [1976] rhythmic theory directly predicts that complexity (i.e., perceivability) of an auditory sequence declines as larger and less constant intervals between pitches are introduced, other work (i.e., Dowling [1978]) has suggested that the processing of contour (i.e., trend changes) and intervals proceeds independently. A follow-up to the current study is planned to disentangle the respective roles of trend reversals and frequency intervals between tones by adjusting data scaling to reflect equal and constant interval changes across different numbers of trend reversals.

Although more research is needed to isolate the exact data properties (interval changes, trend changes, or both) whereby more trend reversals generally tended to result in worse auditory graph comprehension in the current studies, current practices (and current software, etc.) for making graphs from sound have never before considered that trend patterns in the data-to-be-represented may be a critical factor influencing graph comprehension. The current findings suggest that, holding time and frequency scaling factors constant, performance will generally be impacted negatively as the number of trend reversals increases. Performance was at or near the ceiling for the trend-identification task across all data densities when the data had no trend reversals and performance generally declined from ceiling levels as trend reversals were added. The same general pattern of findings was evident in the point-estimation study as well. This does not mean that the introduction of trend reversals will render auditory graphs unusable altogether, and this finding parallels theoretical predictions regarding the comprehension of traditional visual graphs (e.g., Carswell et al. [1993]; Carswell and Ramzy [1997]). It seems that the simplest auditory graphs are characterized by unidirectional, constant frequency change, which elicit Gestalt perceptual good form because of temporal proximity, frequency proximity, and expectations of good continuation of the frequency pattern. The same types of perceptual Gestalts (e.g., spatial proximity, temporal proximity, and good continuation) likely contribute to the automatic perception of patterns in visual graphs as described by Kosslyn [1989], which have also been characterized as *emergent* features or patterns [Sanderson et al. 1989]. It seems plausible that, as data relations become more complex (i.e., as they depart from simple linear increasing or decreasing relations), the comprehension of a graph, either visual or auditory, becomes more difficult. In extreme cases, where data are random and characterized by drastic changes in y-axis values across small changes in x-axis values, no auditory or visual pattern will emerge from graphical representations in any modality, because no pattern is present.

Current practices in auditory graphing generally start by choosing a scaling factor (e.g., 100–2000 Hz), then sonifying the data within that scaling factor. The dilemma regarding the choice of a scaling factor becomes more problematic when data are sonified in real time and a priori maximum and minimum data points are unknown. More research is needed to determine if, as would be predicted by rhythmic theory [M. R. Jones 1976], a proportional scaling change to reduce frequency intervals could be used to alleviate auditory graph comprehensions problems for data with large, rapid changes in a short period of time. Auditory y-axis context in the form of reference tones (see Smith and Walker [2002, 2005]), which was not used in the current study, may also aid in the performance of auditory graphing tasks for data with more frequent trend changes. Furthermore, sonification researchers have yet to look at performance for highly trained listeners. Research has suggested that a brief training session helps naïve auditory graph listeners [e.g., Smith and Walker 2005; Walker and Nees 2005b], and more extensive investigations of training and practice with auditory graphs may show that training interventions also ameliorate some of the detriments in performance observed in the current study as trend reversals in data increased. The data of the current studies, whose design and analyses were not tailored to examine either practice or training, showed only small changes in accuracy comparing early to later trials, but the analyses of performance over time in both studies suggested that participants needed to hear stimuli fewer times by the end of the study to maintain the same level of performance. Interestingly, the general patterns of the decreasing functions for mean times listened were more or less congruent with traditional practice functions (see Newell and Rosenbloom [1981]). Research regarding training with auditory graphs [Walker and Nees 2005b] has suggested that considerable improvements over time might be possible if specific feedback for individual trials is introduced.

The relatively small impact of data density as compared to trend reversals in the current study does not necessarily mean that the timing of data points will have only a small effect on auditory graph comprehension across all reasonable choices of data density per unit of time. The patterns in the current study were isochronous, with regular timing of data points occurring within a given graph. Clearly not all graphical data lend themselves to regular x-axis spacing (e.g., scatterplots). The theories of Martin [1972] and M.R. Jones [1976] both predict that nonisochronous data densities, occurring as a result of data that are irregular on an x-axis value or missing data points, etc., will be problematic for auditory pattern perception. Irregular timing of data, particularly data sets whose timing as auditory graphs follow no predictable hierarchical pattern, will not allow for attention to be focused based on expectancies. Smith and Walker [2002; Smith and Walker 2005], however, have shown that x-axis auditory context—in the form of regular beats or clicks—can help auditory graph users with the temporal organization of data. Such contextual cues may ameliorate the potential problems that could otherwise persist when sonifying nonisochronous data.

Auditory graphs have a great potential to improve data accessibility for blind students and scientists as well as sighted people. This study investigated the role of two important attributes of data—the density and the number of trend reversals. Results suggested a generally more pronounced effect attributable to the role of the number of trend reversals. Furthermore, the current study has suggested that theories of auditory pattern perception (e.g., M. R. Jones [1976]; Martin [1972]) may offer useful insights in understanding how the data in an auditory graph can be presented such that relations are simple and easy to perceive and understand.

ACKNOWLEDGMENTS

This work partially fulfilled thesis requirements for the first author's master's degree. We would like to thank Gregory Corso and Wendy Rogers for insightful comments and feedback on these studies. We also acknowledge the valuable contributions of Dianne Palladino during data collection for this project.

REFERENCES

AMERICAN PSYCHOLOGICAL ASSOCIATION. 2001. Publication Manual of the American Psychological Association (5th ed.). American Psychological Association, Washington, DC.

BONEBRIGHT, T. L., NEES, M. A., CONNERLEY, T. T., AND 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)*, Espoo, Finland. 62–66.

BREGMAN, A. S. 1990. Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, MIT Press, Cambridge, MA.

BREWSTER, S. 1997. Using non-speech sound to overcome information overload. *Displays, 17*, 179–189.

BREWSTER, S. 2002. Overcoming the lack of screen space on mobile computers. *Personal and Ubiquitous Computing, 6*, 3, 188–205.

BREWSTER, S. AND MURRAY, R. 2000. Presenting dynamic information on mobile computers. *Personal Technologies, 4*, 4, 209–212.

BROWN, L. M. AND BREWSTER, S. A. 2003. Drawing by ear: Interpreting sonified line graphs. In *Proceedings of the International Conference on Auditory Display (ICAD2003)*, Boston, MA. 152–156.

BROWN, L. M., BREWSTER, S. A., RAMLOLL, R., BURTON, M., AND RIEDEL, B. 2003. Design guidelines for audio presentation of graphs and tables. In *Proceedings of the International Conference on Auditory Display (ICAD2003)*, Boston, MA. 284–287.

BROWN, L. M., BREWSTER, S., AND RIEDEL, B. 2002. Browsing modes for exploring sonified line graphs. In *Proceedings of the 16th Britich HCI Conference*, London, UK.

BROWN, M. L., NEWSOME, S. L., AND GLINERT, E. P. 1989. An experiment into the use of auditory cues to reduce visual workload. In *Proceedings of the ACM CHI 89 Human Factors in Computing Systems Conference (CHI 89)*. 339–346.

BUTLER, D. L. 1993. Graphics in psychology: Pictures, data, and especially concepts. *Behav. Res. Methods Instruments Comput. 25*, 2, 81–92.

CARPENTER, P. A., AND SHAH, P. 1998. A model of the perceptual and conceptual processes in graph comprehension. *J. Exp. Psych. Appl. 4*, 2, 75–100.

CARSWELL, C. M. AND RAMZY, C. 1997. Graphing small data sets: Should we bother? *Behav. Inform. Technol. 16*, 2, 61–71.

CARSWELL, C. M. 1992. Reading graphs: Interactions of processing requirements and stimulus structure. In *Percepts, Concepts and Categories: The Representation and Processing of Information*, B. Burns, Ed., North-Holland (Elsevier) Boston, MA. 605-645.

CARSWELL, C. M., EMERY, C., AND LONON, A. M. 1993. Stimulus complexity and information integration in the spontaneous interpretation of line graphs. *Appl. Cognitive Psych. 7*, 341–357.

CHILDS, E. 2005. Auditory graphs of real-time data. In *Proceedings of the International Conference on Auditory Display (ICAD 2005)*, Limerick, Ireland.

DEUTSCH, D., AND FEROE, J. 1981. The internal representation of pitch sequences in tonal music. *Psych. Rev. 88*, 6, 503–522.

DOWLING, W. J. 1978. Scale and contour: Two components of a theory of memory for melodies. *Psych. Rev. 85*, 4, 341–354.

DRAKE, C. AND BOTTE, M. C. 1993. Tempo sensitivity in auditory sequences: Evidence for a multiple look model. *Perception Psychophysics 54*, 3, 277–286.

FLOWERS, J. H. AND HAUER, T. A. 1992. The ear's versus the eye's potential to assess characteristics of numeric data: Are we too visuocentric? *Behav. Res. Methods Instruments Comput. 24*, 2, 258–264.

FLOWERS, J. H. AND HAUER, T. A. 1993. "Sound" alternatives to visual graphics for exploratory data analysis. *Behav. Res. Methods Instruments Comput. 25*, 2, 242–249.

FLOWERS, J. H. AND HAUER, T. A. 1995. Musical versus visual graphs: Cross-modal equivalence in perception of time series data. *Human Factors 37*, 3, 553–569.

FLOWERS, J. H., BUHMAN, D. C., AND TURNAGE, K. D. 1997. Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples. *Human Factors 39*, 3, 341–351.

FLOWERS, J. H., BUHMAN, D. C., AND TURNAGE, K. D. 2005. Data sonification from the desktop: Should sound be part of standard data analysis software? *ACM Trans. Appl. Perception 2*, 4, 467–472.

FOLDS, D. J. 2006. The elevation illusion in virtual audio. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, San Francisco, CA. 1576–1580.

FRAISSE, P. 1978. Time and rhythm perception. In *Handbook of Perception. Volume VIII: Perceptual Coding*, M. P. Friedman, Ed., Academic Press, New York. 203–254.

FRAISSE, P. 1982. Rhythm and tempo. In D. Deutsch (Ed.), *The Psychology of Music*. Academic Press, New York. 249–180.

FRIEL, S. N., CURCIO, F. R., AND BRIGHT, G. W. 2001. Making sense of graphs: Critical factors influencing comprehension and instructional applications [Electronic version]. *J. Res. Math. 32*, 2, 124–159.

FRYSINGER, S. P. 2005. A brief history of auditory data representation to the 1980s. In *Proceedings of the International Conference on Auditory Display (ICAD 2005)*, Limerick, Ireland.

GILLAN, D. J. AND LEWIS, R. 1994. A componential model of human interaction with graphs: I. Linear regression modeling. *Human Factors 36*, 3, 419–440.

GILLAN, D. J., WICKENS, C. D., HOLLANDS, J. G., AND CARSWELL, C. M. 1998. Guidelines for presenting quantitative data in HFES publications. *Human Factors 40*, 1, 28–41.

GOBBO, C. 1994. On children's understanding of an economic concept: The role of graphics in evaluation. In W. Schnotz and R.W. Kulhavy (Eds.), *Comprehension of Graphics*. Elsevier Science, Boston, MA. 227–249.

GRUDIN, J. 2001. Partitioning digital worlds: Focal and peripheral awareness in multiple monitor use. In *Proceedings of the 2001 SIGCHI conference on Human Factors in Computing Systems*, Seattle, WA. 458–465.

JONES, M. R. 1976. Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psych. Rev. 83*, 5, 323–355.

JONES, M. R., JOHNSTON, H. M., AND PUENTE, J. 2006. Effects of auditory pattern structure on anticipatory and reactive attending. *Cognitive Psych. 53*, 59–96.

JONES, R. W., AND CARERAS, I. E. 1996. The empirical investigation of factors affecting graphical visualizations. *Behav. Res. Methods Instruments Comput. 28*, 2, 265–269.

KEPPEL, G. AND WICKENS, T. D. 2004. *Design and Analysis: A Researcher's Handbook*, 4th ed., Prentice Hall, Upper Saddle River, NJ.

KOSSLYN, S. M. 1989. Understanding charts and graphs. *Appl. Cognitive Psych. 3*, 3, 185–225.

KRAMER, G. 1994. An introduction to auditory display. In *Auditory Display: Sonification, Audification, and Auditory Interfaces*. G. Kramer Ed., Addison Wesley, Reading, MA. 1–78.

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. International Community for Auditory Display (ICAD), Santa Fe, NM.

KUBOVY, M. 1981. Concurrent pitch segregation and the theory of indispensable attributes. In *Perceptual Organization*, M. Kubovy and J. Pomerantz, Eds., Lawrence Erlbaum, Mahwah, NJ. 55–99.

LEE, J. C. K. AND GERBER, R. 1999. Hong Kong students' perceptions of graphs, charts, and maps. *Scand. J. Educ. Res. 43*, 1, 19–40.

LIU, B., SALVENDY, G., AND KUCZEK, T. 1999. The role of visualization in understanding graphics. *Intern. J. Cognitive Ergon. 3*, 4, 289–305.

MANSUR, D. L., BLATTNER, M. M., AND JOY, K. I. 1985. Sound graphs: A numerical data analysis method for the blind. *J. Med. Syst. 9*, 3, 163–174.

MARTIN, J. G. 1972. Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psych. Rev. 79*, 6, 487–509.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY. n.d. MIT web accessibility guidelines. Retrieved April 10, 2006 from http://web.mit.edu/atic/www/accessibility/developweb.html

MEYER, J. 2000. Performance with tables and graphs: Effects of training and a visual search model. *Ergonomics 43*, 11, 1840–1865.

MEYER, J., SHINAR, D., AND LEISER, D. 1997. Multiple factors that determine performance with tables and graphs. *Human Factors 39*, 2, 268–286.

MOORE, P. J. 1993. Metacognitive processing of diagrams, maps, and graphs. *Learning Instruction 3*, 215–226.

NEES, M. A. AND WALKER, B. N. 2007. Listener, task, and auditory graph: Toward a conceptual model of auditory graph comprehension. In *Proceedings of the ICAD07—Thirteenth International Conference on Auditory Display*, Montreal, Canada (26–29 June). 266–273.

NEWELL, A. AND ROSENBLOOM, P. S. 1981. Mechanisms of skill acquisition and the law of practice. In *Cognitive Skills and Their Acquisition*, J. R. Anderson, Ed., Lawrence Erlbaum Assoc., Mahwah, NJ.

OLIVER, F. 1998. How to present information in graphs and diagrams: Royal Statistical Society.

PALOMAKI, H. 2006. Meanings conveyed by simple auditory rhythms. In *Proceedings of the 12th International Conference on Auditory Display*, London, UK.

PEDEN, B. F. AND HAUSMANN, S. E. 2000. Data graphs in introductory and upper psychology textbooks: A content analysis. *Teaching Psychol. 27*, 2, 93.

PERES, S. C. AND LANE, D. M. 2003. Sonification of statistical graphs. In *Proceedings of the International Conference on Auditory Display (ICAD2003)*, Boston, MA.

PERES, S. C. AND LANE, D. M. 2005. Auditory graphs: The effects of redundant dimensions and divided attention. In *Proceedings of the International Conference on Auditory Display (ICAD 2005)*, Limerick, Ireland. 169–174.

PINKER, S. 1990. A theory of graph comprehension. In R. Freedle Ed., *Artificial Intelligence and the Future of Testing.* Lawrence Erlbaum Associates, Mahwah, NJ. 73–126.

POVEL, D. J. AND ESSENS, P. 1985. Perception of temporal patterns. *Music Perception, 2*, 4, 411–440.

QUESENBERY, W. 1999. Usability interface web accessibility inititative. Retrieved March 11, 2006, from http://www.stcsig.org/usability/newsletter/9910-wai.html

ROTH, P., KAMEL, H., PETRUCCI, L., AND PUN, T. 2002. A comparison of three nonvisual methods for presenting scientific graphs. *J. Visual Impairment Blindness, 96*, 6, 420–428.

SANDERSON, P. M., FLACH, J. M., BUTTIGIEG, M. A., AND CASEY, E. J. 1989. Object displays do not always support better integrated task performance. *Human Factors, 31*, 2, 183–198.

SCHUTZ, H. G. 1961. An evaluation of formats for graphic trend displays–Experiment II. *Human Factors 3*, 99–107.

SHEPARD, R. N. 1982. Structural representation of musical pitch. In *The Psychology of Music*, D. Deutsch, Ed., Academic Press, New York.

SMITH, D. R. AND WALKER, B. N. 2002. Tick-marks, axes, and labels: The effects of adding context to auditory graphs. In *Proceedings of the International Conference on Auditory Display (ICAD2002)*, Kyoto, Japan. 362–367.

SMITH, D. R. AND WALKER, B. N. 2005. Effects of auditory context cues and training on performance of a point estimation sonification task. *Appl. Cognitive Psychol. 19*, 8, 1065–1087.

STOCKMAN, T., HIND, G., AND FRAUENBERGER, C. 2005. Interactive sonification of spreadsheets. In *Proceedings of the International Conference on Auditory Display (ICAD2005)*, Limerick, Ireland. 134–139.

TURNAGE, K. D., BONEBRIGHT, T. L., BUHMAN, D. C., AND FLOWERS, J. H. 1996. The effects of task demands on the equivalence of visual and auditory representations of periodic numerical data. *Behav. Res. Methods Instruments Comput. 28*, 2, 270–274.

TURNBULL, W. W. 1944. Pitch discrimination as a function of tonal duration. *J. Exp. Psychol. 34*, 302–316.

W3C. 2000. Core techniques for Web Accessibility Guidelines 1.0. Retrieved March 12, 2006, from http://www.w3.org/TR/WCAG10-CORE-TECHS/#toc

WALKER, B. N. 2002. Magnitude estimation of conceptual data dimensions for use in sonification. *J. Exp. Psychol. Appl. 8*, 211–221.

WALKER, B. N. 2007. Consistency of magnitude estimations with conceptual data dimensions used for sonification. *Appl. Cognitive Psychol. 21*, 579–599.

WALKER, B. N. AND COTHRAN, J. T. 2003. Sonification Sandbox: A graphical toolkit for auditory graphs. In *Proceedings of the International Conference on Auditory Display (ICAD2003)*, Boston, MA. 161–163.

WALKER, B. N. AND LOWEY, M. 2004. Sonification Sandbox: A graphical toolkit for auditory graphs. In *Proceedings of the Rehabilitation Engineering & Assistive Technology Society of America (RESNA) 27th International Conference*, Orlando, FL.

WALKER, B. N. AND NEES, M. A. 2005a. An agenda for research and development of multimodal graphs. In *Proceedings of the International Conference on Auditory Display (ICAD2005)*, Limerick, Ireland. 428–432.

WALKER, B. N. AND NEES, M. A. 2005b. Conceptual versus perceptual training for auditory graphs. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, Orlando, FL. 1598–1601.

WENZEL, E. M., ARRUDA, M., KISTLER, D. J., AND WIGHTMAN, F. L. 1993. Localization using nonindividualized head-related transfer functions. *J. Acoust. Soc. Amer. 94*, 1, 111–123.

WICKENS, C. D. 2002. Multiple resources and performance prediction. *Theor. Issues Ergon. Sci. 3*, 2, 159–177.

ZACKS, J., LEVY, E., TVERSKY, B., AND SCHIANO, D. 2002. Graphs in print. In *Diagrammatic Representation and Reasoning*, M. Anderson, B. Meyer, and P. Olivier, Eds., Springer, New York. 187–206.

Received June 2007; revised October 2007; accepted November 2007