CONCEPTUAL VERSUS PERCEPTUAL TRAINING FOR AUDITORY GRAPHS

Bruce N. Walker and Michael A. Nees
Sonification Lab
School of Psychology
Georgia Institute of Technology
Atlanta, Georgia, USA

A study examined different types of brief training for performance of a point estimation task with a sonified graph of quantitative data. For a given trial, participants estimated the price of a stock at a randomly selected hour of a 10-hour trading day as displayed by an auditory graph of the stock price. Sixty Georgia Tech undergraduate students completed a pre-test, an experimental training session, and a post-test for the point estimation task. In an extension of Smith and Walker (in press), a highly conceptual, task analysis-derived method of training was examined along with training paradigms that used either practice alone, prompting of correct responses, or feedback for correct answers during the training session. A control group completed a filler task during training. Results indicated that practice with feedback during training produced better post-test scores than the control condition.

INTRODUCTION

Auditory displays of quantitative information have been employed as aids for accomplishing a diverse assortment of tasks in a wide array of settings (Kramer et al., 1999). Sound displays have shown promise when vision is limited, occupied, or impossible, and also as accompaniments to visual displays. While visual depictions of information are pervasive, the portrayal of quantitative information through sound is a relatively unfamiliar phenomenon for novices. Users of auditory displays would benefit from an empirically established process to facilitate learning, yet it remains to be determined what training is best for such displays.

The successful implementation of auditory graphs requires perceptual skill, and research has indicated that perceptual performance can be improved with training (Ahissar, 2001; Aiken & Lau, 1967; Annett, 1966; Annett & Patterson, 1967; Prather, Berry, & Bermudez, 1972; Sidley, Winograd, & Bedarf, 1965; Swets, Millman, Fletcher, & Green, 1962). In many traditional perceptual training paradigms, prompting of correct responses or feedback for correct responses have been employed to improve perceptual acuity. Prompting involves training with a cue that provides the correct response paired to a perceptual stimulus just before or during the presentation of that stimulus. Feedback, on the other hand, requires the trainee to commit to a response before the correct answer is revealed.

The current study compares traditional perceptual training methods to a training paradigm utilized by Smith and Walker (in press) to train auditory graph users. In an auditory graph, quantitative data are represented with sounds, or *sonified*. The visual Y-axis value of a data point is translated to pitch in an auditory graph, and the data point's position along the visual X-axis is represented by time. Smith and Walker investigated the effects of training on a point estimation task with auditory graphs; a group that received a brief training session performed significantly better on the task than did a control group that received no training.

METHOD

Smith and Walker (in press) based their training session on an analysis of the point estimation task. They identified five steps essential to the successful completion of the point estimation task and offered participants strategies for accomplishing each stage of the task. The current study looked at four different types of training for naïve auditory graph users. In addition to a conceptual, task analysis-derived method of training, the perceptual training techniques of prompting and feedback were examined. Finally, a fourth condition involved only practice with the task, while a control group performed a filler task during training. All participants completed a pre-test, an experimental training session (or filler task), and a post-test.

Participants

Sixty Georgia Tech undergraduates participated for partial course credit. The sample included 17 males and 43 females (mean age = 20.3 years, *SD* = 1.61). All reported normal or corrected-to-normal hearing and vision.

Stimuli

Auditory graphs depicted stock price in dollars over the course of a 10-hour trading day (8 a.m. - 6 p.m). Price was represented by discrete tones that changed in pitch as the price changed, while each hour of the trading day corresponded to one second in time. This type of auditory representation is also known as a "tone graph." Discrete tones were presented at the rate of two per second; thus the pitch of each tone represented the price of the stock at a given half-hour of the trading day.

During the pre-test, the auditory graph for all conditions consisted of a pseudo-sinusoidal graph whose values rose for the first 3 hours of the trading day, fell for the next 5 hours, and rose for the last 3 hours. The stock opened and closed at a price of 50 dollars. The pre-test graph was presented without context such that participants heard only the sonified data points. During the post-test phase, participants in all conditions again heard the same auditory graph, but the post-test graph was enriched with both X-axis context and Y-axis context.

Auditory graphs were given Y-axis (price) context using a dynamic Y reference tone that bracketed sets of consecutive ascending or descending tones. The bracketing reference was a beeping tone displayed concurrently with the actual discrete data points of the auditory graph to create the sound equivalent of visual gridlines for Y-axis values. The tone represented the stock's maximum price of the day when the true price of the stock was ascending and indicated the stock's lowest price of the day when the true price was descending. Time context was provided by the addition of a rhythmic beat on every hour in the display. The bracketing Y reference and the X-axis click track have been shown to provide helpful auditory context for point estimation tasks with auditory graphs (Smith & Walker, 2002, 2004, in press).

A different auditory graph was used where applicable during the training phase of the study; the values of the stock in the training graph assumed an inverse sinusoidal shape. The different training graph was employed to prevent simple memorization of stock price values from confounding the post-test data. The training graph used both X and Y-axis context.

Procedure and Task

Participants were randomly assigned to one of five experimental training conditions. The pre-test was the same for all participants. For the 11 pre-test trials, participants were asked to identify the price of the stock for each hour of the trading day in a randomly selected order. Participants were told that the opening price of the stock was 50 dollars, and a single trial began with a visual 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. Participants were allowed to listen to the auditory graph as many times as needed before responding, and the next trial began after a response had been made using the computer keyboard. Participants were given a two minute break following the pre-test.

Participants then completed the experimental training manipulation for their assigned conditions. The training sessions were designed to be self-paced and approximately equivalent in duration for each of the five conditions, and participants received a two minute break following completion of training.

The post-test segment of the study began immediately following the final break. All participants were given the same task from the pre-test, where the price of the stock was estimated for each hour of the trading day over 11 trials of randomly selected hours. The post-test auditory graph, which did employ context, was the same for all conditions.

Experimental Conditions

Five training conditions were examined; the conditions varied with respect to the type of material presented during the training session between pre- and post-tests.

Group 1. This control condition experienced a filler task during training.

Group 2. This group experienced only practice with auditory graphs as training. Participants completed 22 trials (two complete sets of each hour in the trading day) of the point estimation task during training. This condition examined the potential facilitative effects of mere exposure to additional trials of the point estimation task with auditory graphs.

Group 3. This condition completed 22 trials of the task with the aid of feedback regarding performance. Participants listened to the training auditory graph and responded; disclosure of the correct response followed.

Group 4. This condition completed 22 trials of the task with the guidance of a visual prompt during training. Participants heard the training graph, and the visual display concurrently presented text indicating the time of day (X-axis value) and stock price (Y-axis value).

Group 5. This condition experienced the same conceptual training procedures as described in Smith and Walker (2004; in press). An interactive presentation with a voice-over and visual aids explained sonification and auditory graphs.

RESULTS

The dependent variable was the root mean squared (RMS) error (in dollars) of participants' responses to the point estimation trials, and RMS error scores were analyzed with a series of planned analysis of covariance (ANCOVA) contrasts. Participant pre-test scores were entered into the analyses as the covariate. The correlation between pre-test and post-test RMS error scores was significant, $r = .35$, $p < .05$, and statistical assumptions of ANCOVA were verified.

Table 1. Adjusted mean post-test RMS error scores.

Group	Type of training	ANCOVA adjusted mean RMS error (in dollars)	*S.E.*
1	Filler task (control)	20.43	3.03
2	Practice only	17.31	2.94
3	Practice with feedback	10.00	2.94
4	Practice with a visual prompt	16.17	2.93
5	Conceptual training	14.89	2.92

Table 1 displays the adjusted means and standard errors for each of the five treatment conditions. Four planned comparisons examined each of the four training groups (Groups 2, 3, 4, and 5) in direct comparisons with the filler control group (Group 1). Practice alone (Group 2) was not significantly different from the no training condition, $F\ (1,54) = .53$, $p > .05$. Practice with feedback resulted in significantly better post-test performance (lower RMS error) than no training, $F\ (1,54) = 5.91$, $p < .05$. For the contrast between Groups 1 and 3, a large effect size (see Cohen, 1992) of group membership on RMS post-test scores was present, Cohen's $d = 1.01$. Both prompting (Group 4) and conceptual training (Group 5) failed to exhibit a significant mean difference from the filler task for post-test RMS error, $F\ (1, 54) = 1.01$, $p > .05$; $F\ (1,54) = 1.75$, $p > .05$, respectively.

Figures 1-5 graph the unadjusted raw stock price response data for each of the five groups across the 11 trials of the post-test. The bold lines represent the correct answers on the post-test, and greater deviations from the bold lines represent more error in the response patterns of participants within the condition which is reflected in larger RMS error values.

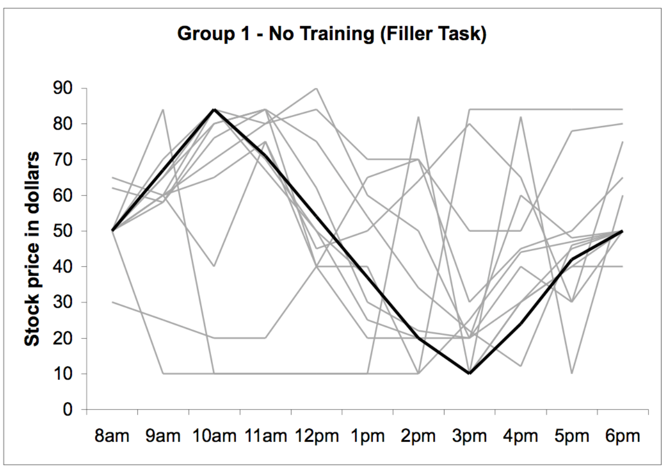

Figure 1. Unadjusted raw data for Group 1 over the 11 post-test trials (gray lines). The bold black line indicates the pattern of correct responses, and deviations from the bold line reflect response error.

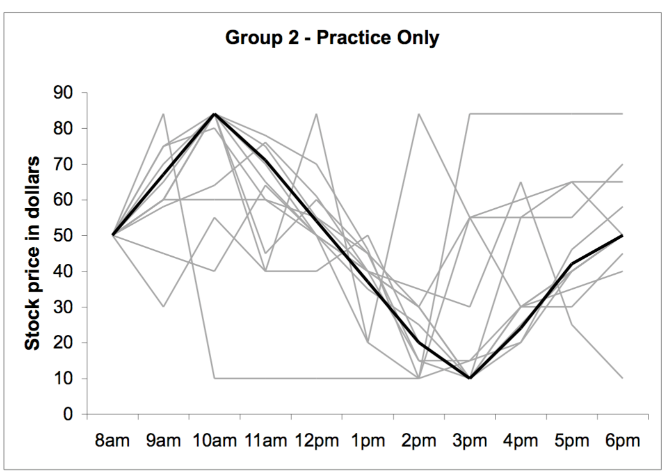

Figure 2. Unadjusted raw data for Group 2.

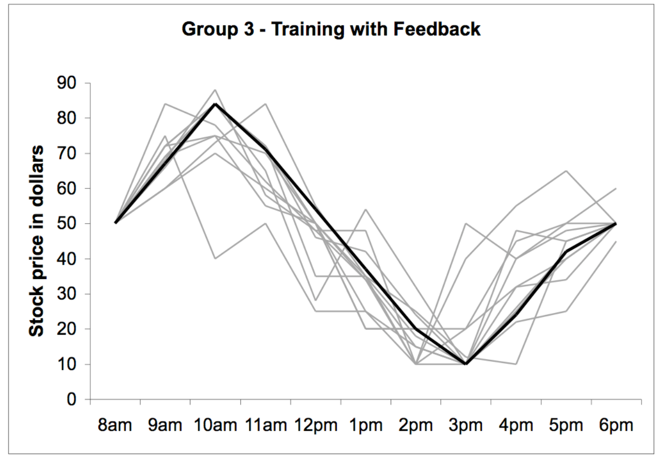

Figure 3. Unadjusted raw data for Group 3.

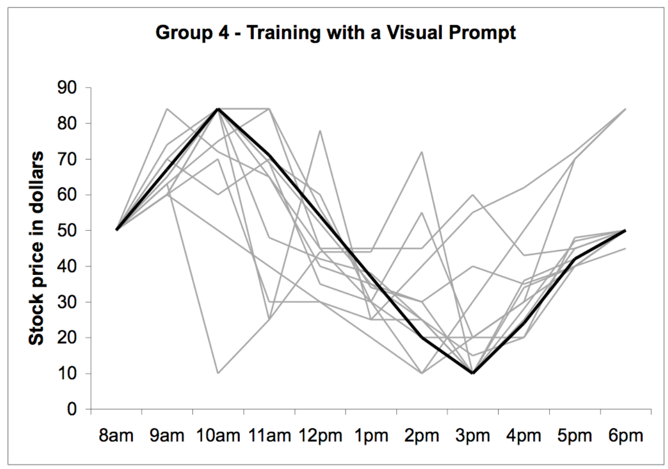

Figure 4. Unadjusted raw data for Group 4.

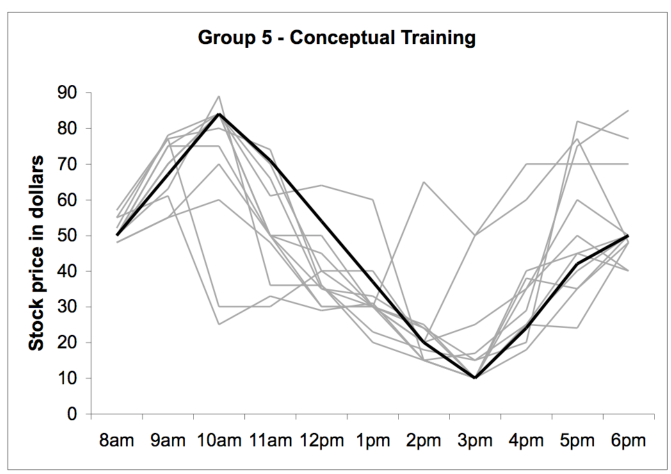

Figure 5. Unadjusted raw data for Group 5.

DISCUSSION

The current study examined the effects of qualitatively different types of training for a point estimation task with auditory graphs. Although Group 5 was trained to successfully accomplish each of the individual stages of the point estimation task, Group 3 was the only condition to exhibit significantly superior post-test performance over the control group. Group 3 was trained with feedback, a technique that has been employed in previous perceptual learning research. Intuitively, this pattern of results indicates that the perceptual elements of the task, namely judging the values corresponding to pitches of tones, may be more difficult to accomplish than other aspects of the task and deserve special attention in an ideal training session. Even in the initial stage of exposure to auditory graphs, the benefit of providing participants with a conceptual tutorial for auditory graphs was overshadowed by the naïve users' need to simply practice the task with feedback regarding the correctness of their responses.

The current study failed to replicate the results of Smith and Walker (in press) regarding conceptual, task-analysis derived training for auditory graphs. ANCOVA adjusted mean results suggest a trend whereby Smith and Walker's (in press) conceptual training condition exhibited a non-significant mean difference in post-test RMS error as compared to no training. The current study employed a smaller sample size and more training conditions than Smith and Walker; thus, the present analysis may not have been powerful enough to detect an effect. Alternatively, elaborate conceptual explanations may be unnecessary in training for auditory graph usage; feedback during training may be sufficient even in the early stages of learning. Further research is required (and ongoing) to resolve these discrepant findings.

An unanticipated but noteworthy phenomenon can be observed in Figures 1 and 2 (the raw data for the filler and practice only conditions). A select few participants seemed to be unable to distinguish Y-axis contextual reference tones from the actual sonified data points despite the fact that instruments of different timbres were used for data and context. The responses for these participants perseverate on the maximum and minimum values that were presented as beeping context cues. Future research on auditory graphs should examine methods of promoting auditory gestalt perceptual grouping such that data is always salient from context.

The successful deployment of auditory graphs will require effective and efficient training programs. While conceptual training has been shown to be effective in past research (Smith & Walker, in press), other approaches such as practice with feedback may be even more effective. The training session in the current study was a single, brief session, and accuracy (as measured by RMS error) was the only outcome measured. Further research should examine training over multiple sessions in order to determine an optimal training regime for speed and accuracy. In all likelihood, the type of training that best facilitates improved performance will change across stages of skill acquisition. Ultimately, researchers will need to examine transfer of training across tasks (point comparison, trend evaluation, extrapolation, etc.) with auditory graphs, and a successful training regime for auditory graphs should be adapted for other types of auditory displays.

REFERENCES

Ahissar, M. (2001). Perceptual training: A tool for both modifying the brain and exploring it. *Proceeding of the National Academy of Sciences, 98*(21), 11842-11843.

Aiken, E. G., & Lau, A. W. (1967). Response prompting and response confirmation: A review of recent literature. *Psychological Bulletin, 68*(5), 330-341.

Annett, J. (1966). Training for perceptual skills. *Ergonomics, 9*(6), 459-468.

Annett, J., & Patterson, L. (1967). Training for auditory detection. *Acta Psychologica, 27*, 420-426.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.

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).

Prather, D. C., Berry, G. A., & Bermudez, J. M. (1972). The effect of prompting and feedback on performance during learning, stress, and transfer of a perceptual skill. *Proceedings of the Annual Convention of the American Psychological Association, 7*(2), 643-644.

Sidley, N. A., Winograd, E., & Bedarf, E. W. (1965). Learning to identify complex sounds: Prompting versus confirmation. *Journal of the Acoustical Society of America, 38*, 1050-1052.

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

Smith, D. R., & Walker, B. N. (2004). Effects of training and auditory context on performance of a point estimation sonification task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting (HFES2004)* (pp. 1828-1831), New Orleans, LA.

Smith, D. R., & Walker, B. N. (in press). Effects of auditory context cues and training on performance of a point estimation sonification task. *Journal of Applied Cognitive Psychology.*

Swets, J. A., Millman, S. H., Fletcher, W. E., & Green, D. M. (1962). Learning to identify nonverbal sounds: An application of a computer as a teaching machine. *Journal of the Acoustical Society of America, 37(7)*, 928-935.