EFFECTS OF TRAINING AND AUDITORY CONTEXT ON PERFORMANCE OF A POINT ESTIMATION SONIFICATION TASK

Daniel R. Smith
Department of Behavioral Sciences & Leadership
United States Military Academy
West Point, NY 10996
daniel.smith@usma.edu

Bruce N. Walker
Sonification Lab, School of Psychology
Georgia Institute of Technology,
Atlanta, GA 30332-0170
bruce.walker@psych.gatech.edu

Abstract

Research on auditory graphs has investigated mappings, scalings, and polarities (Walker, 2002), as well as the addition of some contextual design features (Bonebright, Nees, Connerley, & McCain, 2001; Flowers, Buhman, & Turnage, 1997), in order to improve performance. However, little has been done to quantify the performance effects of such features, or to investigate effects of training in specific sonification tasks such as point estimation. Smith and Walker (2002) took a step towards quantifying and comparing the effects of adding several contextual design features. Presented here are selected results from a comprehensive follow-on study comparing effects of adding auditory context, either with or without training. The overall results indicate that some kinds of auditory context improved performance, while others did not. Training improved performance, and an interaction was discovered between type of auditory context and type of training (Smith, 2003). Implications are discussed.

INTRODUCTION

The most common tools and techniques for displays of quantitative information are almost entirely visual. However, such unimodal visual displays can be problematic for a number of reasons (Kramer et al., 1999), but especially because of the difficulty that occurs when the operator is either unable to *look at* or unable to *see* a visual display, or when the visual system is occupied with another task (Cohen, 1994; Wickens & Liu, 1988). Sonification, the use of an auditory display to convey quantitative information, can provide flexibility for the design of displays, but questions remain about how to increase performance to levels that enable its use in more complex systems (Johannsen, 2002; Marila, 2002).

The defining feature of an auditory graph (as opposed to a more general sonification) is the inclusion of auditory context elements intended to assist the user in a more accurate perception of the data relations. Given these elements (e.g., auditory equivalents of tick-marks, axes, and labels), auditory graphs become analogous to visual graphs and should yield more accurate perceptions of the data (Walker, 2002). In addition, considering the relative inexperience of the general population in the use of sonification and auditory graphs, it seems probable that training could also improve performance (Kramer et al., 1999).

Although work has been done to investigate how designers might add context to a display, little has been done to quantify and compare the performance effects of such features - or to investigate effects of training in specific sonification tasks such as point estimation. Therefore, we explicitly tested the effects of added auditory context, with and without training focused on the cognitive and perceptual skills hypothesized to be operative in the perception of auditory graphs.

METHOD

Participants

A total of 150 undergraduates from Georgia Tech and the United States Military Academy participated for course credit. There were 82 males and 68 females, with a mean age of 20.1 years (*SD* = 2.45). All reported normal or corrected to normal vision and hearing, and provided demographic details about age, sex, handedness, and number of years of musical training.

Apparatus

Instructions and visual stimuli appeared on a 17-in. (43.2 cm) Apple Macintosh studio display set to a resolution of 1024 x 768 pixels, and viewed from a distance of 24 in. (61 cm). Auditory stimuli were presented via headphones.

Design

There were two main factors of interest: training (present or absent), and the presence of auditory context cues, analogous to the axes and tick marks in a visual graph. This was subdivided into the presence or absence of “x-axis” context (auditory cues that help the listener mark the passage of time) and the presence or absence of “y-axis” context (cues that help the listener determine the value of the data at a given point in time). These variables were manipulated between subjects in a fully factorial 2 x 2 x 2 design consisting of training, x-context and y-context.

Procedure

There were two blocks of trials. The first served as a measure of baseline performance, and was the same for all participants. The second block was the key experimental block, wherein differences between the conditions would be assessed. The training variable was operationalized by the presence or absence of a training program between the two blocks, and the context variables were manifested by the presence or absence of auditory context in the second block only.

On a given trial (regardless of block), participants listened to a 10-second auditory graph representing the price of a stock as it fluctuated through a 10-hour trading day. In the first block all participants heard the same sounds: a basic auditory graph mapping stock price onto pitch, with each hour lasting one second. This simple sound structure and lack of training meant performance relied entirely on the intuitions of the novice participant. Following the presentation of the auditory graph, participants estimated the stock price at a specific hour of the day. In total, 11 such trials were completed per block, with the times of day randomly ordered for each participant. Following Block 1, half of the participants took part in a training program aimed at improving their understanding and performance on the point estimation task. The training broke the task into its component supporting tasks, as is often recommended (Adams, 1987; Proctor & Dutta, 1995; Quinones & Ehrenstein, 1996; Salvendy, 1997; Smith & Walker, 2002). The nature of each supporting task was explained, and strategies practiced for the completion of the task in the presence or absence of each of the applicable design features for that group. The other half of the participants (i.e., those in the No Training condition) received no formal training in auditory graph comprehension between the blocks. Instead, they completed a reading comprehension filler task.

In Block 2, participants completed the listening task again. Participants in the No Context condition heard the same auditory graph heard in Block 1. Participants in the X-context condition experienced the graph with the addition of x-axis context. This was created by the insertion of audible clicks representing the passing of the hours in the 10-hour trading day (1 click/sec). Participants in the Y-context condition experienced the graph with the addition of y-axis context that was created by the addition of a dynamic tone used to indicate the minimum and maximum range in the data, for calibration of value (y-axis) judgments. When the price of the stock was rising the pitch of the reference tone corresponded to the highest price of the day. When the price was falling, the pitch of the reference tone changed to match the lowest price. Finally, participants in the Dual-context condition experienced the graph with the combination of x-axis context (clicks), and y-axis context (the dynamic min/max reference tone). Again, in Block 2 all participants completed the same 11 value estimation questions following presentation of their respective plain or enhanced auditory graphs.

RESULTS

In both blocks the dependent variable was the root mean squared (RMS) error (in dollars) with which a participant reported the 11 queried data values

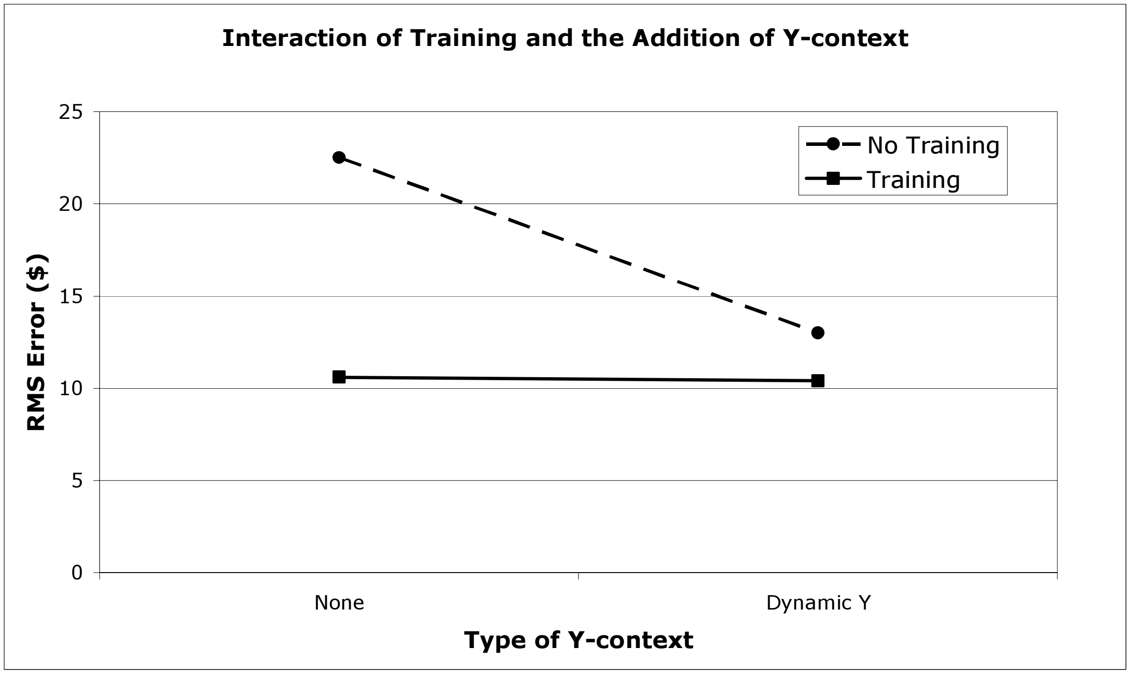

Figure 1. Graph of the results for RMS (in dollars). Overall, participants with training (bottom solid line) performed better than participants without training (dashed top line). Note the interaction between Training and Type of Y-axis context, indicating that the addition of dynamic y-axis context can have an effect on performance that is similar to training.

represented in the display. A three-way ANCOVA was conducted on Block 2 scores, using type of x and y-axis context, and level of training as the between-subjects independent variables; Block 1 score was the covariate.

The results discussed here are shown in Figure 1. There was a significant main effect of training, reflecting that listeners answered with smaller errors if they received training $F(1,150) = 10.405$, $p < .01$. There was also a significant main effect of type of y-axis context, demonstrating that listeners answered with smaller errors if they had the dynamic (min/max-price) y-axis reference tone $F(1,150) = 4.92$, $p < .05$. Finally, there was a significant training by y-axis context interaction reflecting that the effect of the auditory context was dependant upon level of training $F(1,150) = 4.774$, $p < .05$. This interaction, combined with post-hoc comparisons, indicate that when groups received training, adding y-axis context did not produce significant shifts in the means. On the other hand, if groups did not receive training, adding y-axis context did improve performance (see Figure 1).

DISCUSSION

The results confirm theoretical predictions expecting significant main effects of training and at least one of the types of context. The interaction between amount of training and added context is also as predicted.

Training results in overall better performance (lower error; solid line in Figure 1) since the knowledge participants acquire during their training session plays an important role in calibrating the listener's scaling of both x and y-axis dimensions. This information is not available to the untrained participants. As a result, in the untrained condition (dashed line), we see a larger error, overall.

Overall, the addition of y-axis context also produces a large improvement in performance (left versus right sides of Figure 1). In terms of the interaction, in the trained conditions we find relatively little differences between contextual settings. This is partially because of floor effects, since error gets quite low with either training or y-context; both together have limited added effect in this experiment. In the untrained group, however, there is a clear effect of adding a dynamic reference tone. This is because in that condition the context sound is the only means by which scaling information is communicated to the listener.

We conclude that both context and training can produce similar performance effects and ultimately yield the same level of performance. Furthermore, it seems that although context affects the performance of untrained users, it does not produce significant differences in the performance of trained users. This may be because training and context provide overlapping information - or it may be that the particular training programs in this study were not effective enough in enhancing participants' abilities to use the context provided.

Overall, these results provide empirical evidence that: (1) training programs focused on the cognitive and perceptual skills used in sonification interpretation are an effective means of increasing performance; (2) that the effectiveness of added auditory context depends on the level of training given to the user; and (3) that the types of context chosen for the display may alter the effectiveness of training in several of the specific tasks required in the performance of a point estimation task.

Just as a single line on a page does not constitute a visual graph (it lack axes, tick marks, and so on), a single variable pitch does not necessarily constitute an auditory graph. The present results demonstrate this point clearly, and also point out the benefits of not only practice, but also training that teaches a listener to make use of all of the sounds (including auditory context) that may be present. Ongoing research will continue to determine the best ways of communicating numerical information with auditory graphs.

REFERENCES

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. Psychological Bulletin, 101(1), 41-74.

Bonebright, T. L., Nees, M. A., Connerley, T. T., & McCain, G. R. (2001, July 29 - August 1). Testing the effectiveness of sonified graphs for education: A programmatic research project. Proceedings of the International Conference on Auditory Display (pp. 62 - 66), Espoo, Finland.

Cohen, J. (1994). Monitoring background activities. In G. Kramer (Ed.), Auditory display: Sonification, audification, and auditory interfaces (pp. 449-531). Reading, MA: Addison-Wesley.

Flowers, J. H., Buhman, D. C., & Turnage, K. D. (1997). Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples. Human Factors, 39(3)(September), 341-351.

Johannsen, G. (2002, July). Auditory display of directions and states for mobile systems. Proceedings of the International Conference on Auditory Display (pp. 98-103), Kyoto, Japan.

Kramer, G., Walker, B., Bonebright, T., Cook, P., Flowers, J., Miner, N., et al. (1999). The sonification report: Status of the field & research agenda: The International Community for Auditory Display.

Marila, J. (2002, July). Experimental comparison of complex and simple sounds in menu and hierarchy sonification. Proceedings of the International Conference on Auditory Display (pp. 104-108), Kyoto, Japan.

Proctor, R. W., & Dutta, A. (1995). Skill acquisition and human performance. Thousand Oaks, Calif.: Sage Publications.

Quinones, M. A., & Ehrenstein, A. (1996). Training for a rapidly changing workplace: Applications of psychological research. Washington D.C.: American Psychological Association.

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

Smith, D. R. (2003). Effects of training and context on human performance in a point estimation sonification task. Unpublished Master's thesis, Georgia Institute of Technology, Atlanta, Georgia.

Smith, D. R., & Walker, B. N. (2002, July). Tick-marks, axes, and labels: The effects of adding context to auditory graphs. Proceedings of the International Conference on Auditory Display (pp. 362-368), Kyoto, Japan.

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

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