Mappings and Metaphors in Auditory Displays: An Experimental Assessment.

Bruce N. Walker
Psychology Department, Rice University
P.O. Box 1892,
Houston, Texas 77251
Phone: (713) 527-8101
Fax: (713 285-5221
Email: walkerb@Rice.edu

Gregory Kramer
Clarity/Santa Fe Institute
310 NW Brynwood Lane
Portland, OR 97229
503-292-8550
Fax: 503-292-4982
kramer@listen.com

ABSTRACT

Auditory displays are becoming more and more common, but there are still no general guidelines for mapping data dimensions (e.g., temperature) onto display dimensions (e.g., pitch). This paper presents experimental research on different mappings and metaphors in a generic process-control task environment, with reaction time and accuracy as dependent measures. It is hoped that this area of investigation will lead to the development of mapping guidelines applicable to auditory displays in a wide range of task domains.

KEYWORDS

Mapping, metaphor, guidelines.

1. Introduction

Sound has been used in human-system interfaces for many years (e.g., Patterson, 1982; Pollack & Ficks, 1954). Until recently, however, the majority of these audible cues have been simple warning sounds. Sonification, wherein data is represented directly by one of many possible sound attributes, or dimensions, is rapidly maturing (cf., Kramer, 1994a) but is still at the technical and conceptual stage that visual display was a few decades ago. More and more applications use sound to convey information, but just as was the case with early visual displays there are currently no standards, and interface designers have usually implemented what sounds "good" to them. In addition, few of the designers have tested their auditory displays within a rigid experimental setting.

The principles for designing effective visual displays are quite generic, in that they apply to displaying all sorts of information, across a wide variety of task domains (e.g., Shneiderman, 1992; Tufte, 1990). We are now investigating whether generalizable guidelines for auditory displays can be determined as well. In particular, we are examining metaphors employed in the mapping of data dimensions (e.g., temperature) onto display dimensions (e.g., pitch). For example, representing a rising temperature with a rising pitch seems a "natural" choice; it makes intuitive sense. Our mental model of the data space seems to correlate well with the display space. But are there other such natural mappings?

An important consideration is whether a particular mapping choice has a measurable effect on the performance of a task which relies on the auditory display. Thus, are there better ways to represent temperature? Would another mapping produce faster or more accurate responses? Are some mappings more pleasing or easier to understand? We are especially concerned with representations of common data dimensions that may appear in a wide variety of auditory displays. However, along with the usual temperature, pressure, size, cost, and rate, we are also very interested in how best to display more subjective and affective variables such as "value", "goodness", "beauty", "risk" and so on (Kramer, 1994b).

In addition to the decision of which data dimension to represent by which auditory feature, the *direction* of the mapping is often critical. The temperature-to-pitch mapping seems natural only as long as rising pitch signals a rise in temperature. We still do not know whether the "inverse" mapping (i.e., rising pitch signals a *drop* in temperature) would actually affect performance on a task that relied on that auditory display. Some mappings are based on very common or "dead" metaphors (Lakoff & Johnson, 1980), and we can intuitively decide which direction makes more sense to us. There are many cases, though, where it is difficult to predict which direction of a mapping will produce superior results. If "voltage" were mapped onto "richness" (number of harmonics, for example), should an increase in voltage be represented by an increase or a decrease in the number of harmonics? To really find out which direction of this mapping is more effective, we need a performance measure based on a task that requires the auditory display.

2. Procedure

In order to measure performance in a task setting, yet still pursue generalizable, task-independent mapping results, we have developed a generic process control (a "crystal" factory) as our experimental environment (cf., Gaver, Smith & O'Shea, 1991). This way we can include virtually any type of data dimension (including affective variables), and have complete control over how the variables interact and how they are displayed.

Participants listen to the auditory display via headphones in a sound-attenuated room, and they make responses using a response box consisting of rows of large buttons. Each participant receives a basic description of the Crystal Factory, and is trained to associate each data dimension (e.g., the pressure of the crystal formation process) with a dimension of the auditory display (e.g., "brightness" of a sound). This training involves both a verbal description and auditory practice.

The actual environment involves four variables at this time, each of which controls one aspect of the audio output. The data values all remain at their starting points for several seconds, and then one of the variables increases or decreases. The listener hears this as a period of steady state in the factory process followed by a change in one of the process parameters. This is explained to the listener as a period of "normal operations" followed by "something going wrong". In order to preserve the crystal quality, he or she is be required to make an appropriate control action as quickly as possible, using the labeled response buttons. For example, if the temperature drops (perhaps represented by an increased loudness of the sound), then the correct response would be to press the "Heater On" button (see Fitch & Kramer, 1994, for a similar design).

Subjects all hear the same actual sounds, but are required to make different responses depending on their training condition. There are several different trial types, varying the starting values of the variables, and the variable that changes. Each trial type is repeated several times within a block of trials, and each participant completes several blocks of trials. The independent variables include the particular mapping that the listener had been trained to hear, and the actual variable that changes on a given trial. The performance measures include response time (RT) and accuracy. In addition, after each trial the participant is asked to say which parameter of the process changed, to ensure that he or she is paying attention to the metaphor, and not simply mapping the auditory display parameter directly to the response button.

3. Predictions

This work is still underway, due to technical difficulties encountered while setting up the study. Full details of the results will be presented at ICAD. However, we do present some definite predictions. Mappings that are based on stronger or more natural metaphors will result in faster and more accurate control reactions. They should also be learned faster, which will lead to a larger improvement in performance across the blocks of the experiment. For some mappings there will also be a particular direction that results in better performance (e.g., rising temperature mapped to rising, as opposed to falling pitch). These results should complement the findings in the area of stimulus-response compatibility (e.g., Proctor & Reeve, 1990) and cross-modality matching (Melara & O'Brien, 1990; Walker & Ehrenstein, 1996).

4. Implications

It is likely that a number of the most "successful" mappings will be the ones that have most often been used in auditory displays. However, we expect to discover other good mappings, and in particular we will try to display variables that have great possibilities, but have not often been represented with sound. The strong emotive power of music (cf., Révész, 1954) suggests that affective variables are perfect examples of information that may be difficult to describe with words or pictures, but will be easily recognized with sound.

This research is a first step in attempting to quantitatively compare different auditory display setups. We are careful to note that the design of an effective auditory display will always require practice and good judgment. However, the extension of the present research may help to identify guidelines for representing data with sound, which will hopefully apply across a wide range of task domains.

5. ICAD Presentation

For the presentation at the International Conference on Auditory Display, we plan to discuss the actual results that we obtain. We will present several examples of the actual sounds used in the experiment, so that the audience can get a first-hand sense of our mapping conditions. In particular, we will play examples of the most natural mappings, as well as some of the more ambiguous mappings, and discuss how performance varied in the different conditions. Further implications will be addressed, and we hope to generate a lively discussion about the possibilities of general guidelines for the use of mappings and metaphors in auditory displays.

6. References

Fitch, T., & Kramer, G. (1994). Sonifying the body electric: Superiority of an auditory display over a visual display in a complex, multivariate system. In G. Kramer (Ed.), *Auditory display: Sonification, audification, and auditory interfaces*. Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XVIII. Reading, MA: Addison Wesley. 307-325.

Gaver, W., Smith, R. B., & O'Shea, T. (1991). Effective sounds in complex systems: The ARKola simulation. Proceedings of CHI'91, held April 28-May 2, 1991, New Orleans. Reading, MA: ACM Press/Addison Wesley.

Kramer, G. (1994a). An introduction to auditory display. In G. Kramer (Ed.), *Auditory display: Sonification, audification, and auditory interfaces*. Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XVIII. Reading, MA: Addison Wesley. 1-77.

Kramer, G. (1994b). Some organizing principles for representing data with sound. In G. Kramer (Ed.), *Auditory display: Sonification, audification, and auditory interfaces*. Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XVIII. Reading, MA: Addison Wesley.185-221.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Melara, R. D., & O'Brien, T.P. (1990). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General, 116(4)*, 323-336.

Patterson, R.D. (1982). *Guidelines for auditory warning systems on civil aircraft.* Paper No. 82017, Civil Aviation Authority, London.

Pollack, I., & Ficks, L. (1954). Information of elementary multidimensional auditory displays. *Journal of the Acoustical Society of America, 26*, 155-158.

Proctor, R.W., & Reeve, T.G. (Eds.) (1990). *Stimulus-response compatibility: An integrated perspective*. Amsterdam: North Holland.

Révész, G. (1954). *Introduction to the psychology of music*. Norman, OK: University of Oklahoma Press.

Shneiderman, B. (1992). Designing the user interface: Strategies for effective human-computer interaction (2e.). Reading, MA: Addison Wesley.

Tufte, E. (1990). *Envisioning information*. Graphics Press.

Walker, B.N., & Ehrenstein, A. (1996). Cross-dimensional compatibility effects with dynamic auditory stimuli. Paper presented at the Southwest Psychological Association 42nd Annual Convention. Houston, April 4-6, 1996.