The ear's versus the eye's potential to assess characteristics of numeric data: Are we too visuocentric?

JOHN H. FLOWERS and TERRY A. HAUER
University of Nebraska, Lincoln, Nebraska

A single experiment studied how effectively information about the central tendency, variability, and shape of numeric data distributions could be conveyed to statistically knowledgeable subjects. The data were summarized by visual histograms, auditory histograms that coded numeric value as pitch on the musical scale, and five-note auditory analogues of a box-whisker display that coded the minimum, quartile, and maximum scores as musical notes. Regression and multidimensional scaling analysis of judgments of dissimilarity between distributions showed that auditory depiction provides a highly effective means of conveying information about distributional characteristics. Auditory depiction may be a useful alternative to traditional visual graphics.

Modern computer technology has made possible the development of novel forms of graphical data depiction and display, which are being used increasingly in interactive software designed for exploratory data analysis. Visual graphics, including dynamic displays and various systems for plotting multidimensional data, are proving especially useful for research in the behavioral and social sciences. On the other hand, far less attention has been paid to the development of systems that use sound as a data descriptor, even though software developers are increasingly making use of sound effects as a signaling medium and as an enhancement to "presentation" software. Although the potential of using sound encoding of numeric data (sonification) has been recognized by a few researchers in computer science and engineering (Blattner, Greenberg, & Kamegai, 1990; Bly, 1982), and some development of specific hardware and software tools for data sonification is beginning to take place (Scalletti & Craig, 1991), basic psychological research into the similarity and differences between the auditory and visual systems' abilities to provide information about data characteristics remains relatively unexplored. It is our position that comparisons of auditory and visual graphics as data descriptors presents an interesting research domain for psychologists concerned with cross-modal and multimodal perception. Basic research in this area is essential for designing efficient enhancements to traditional graphic displays for normal users as well as data display and analysis systems specifically tailored for the visually impaired.

In the present investigation, we studied the ability of statistically knowledgeable subjects (advanced graduate students in experimental and social psychology) to perceptually assess differences in characteristics of numeric data samples that varied in central tendency, variability, and shape (kurtosis and skewness). In separate trial blocks, these data distributions were depicted as traditional visual histograms, and as two varieties of auditory displays: auditory histograms and auditory quartile displays.

In most previous experimental comparisons of the efficiency of auditory and visual presentation of data, some form of identification task involving accuracy of recognition of a previously presented display has been the index of performance (e.g., Bly, 1982). Such tasks involve a substantial memory component that is not encountered, for example, in tasks that simply require judgment or description of differences between two or more data samples. It is our view that comparative assessment of data is more typical of the actual use of graphical displays in exploratory data analysis than is the need to determine that a sample is absolutely identical to one encountered previously. Furthermore, recognition tasks do not provide a straightforward means of assessing possible differences in the *perceptual structure* of a set of stimuli when it is presented in different sensory modalities. In contrast, data from stimulus comparison tasks, when subjected to regression and multidimensional scaling analyses, can be quite informative about the relative salience of stimulus properties in different modalities (see, e.g., Garbin, 1988). To provide an overview of how the auditory and visual data depiction modes convey information about each of several characteristics of distributional data (e.g., central tendency, variability, shape), as well as to assess the "overall effectiveness" of each mode, we chose to use a task that required rating of dissimilarity between pairs of data displays.

METHOD

Subjects
Twelve advanced graduate students in psychology served in three 45-min sessions and received $10 for their time. All subjects had

---

Correspondence should be addressed to John H. Flowers, Department of Psychology, 209 Burnett, University of Nebraska, Lincoln, NE 68588-0308.

258

normal vision and hearing, and all had been actively involved in quantitative data collection in the behavioral sciences.

Stimulus Materials

Twelve samples of 50 numeric values were computer generated to create distributions varying in central tendency and shape, as follows. Six of the distributions were synthesized to approximate samples from an essentially normal Gaussian distribution (by summing 10 values from a rectangular random number generator). Three of the distributions were synthesized directly from a rectangular random number generator, thus producing three highly platykurtic samples. Three other distributions were synthesized to produce positively skewed distributions (by adding approximately normal distributions to the low end of a rectangular distribution). The means of these distributions were adjusted to fall into three groupings (about 44, 49, and 55), with two "normal," one platykurtic, and one skewed distribution in each grouping. The standard deviations were constrained to be relatively constant during distribution synthesis; the actual values ranged from 7.3 to 9.8. Stimulus parameters for each of the 12 distribution "stimuli" are summarized in Table 1.

During each experimental session, the subjects judged the dissimilarity of 132 pairs of distributions, depicted in one of three display modes. *Auditory histograms* displayed each distribution by coding numeric value as pitch of the diatonic musical scale. The numeric values of each of the scores in the first distribution were scaled to fit in a single octave. This display began with the lowest score category or "bin" represented by middle C, producing a 13-category "histogram" in which the pitch of the note represented the numeric value category, and in which the number of times a note was repeated represented the frequency of scores in that category. The notes for the second (comparison) distribution were also scaled to the lowest score and range of the *first* (standard) distribution, and thus could begin above or below middle C and have a range of more than or less than a single octave, depending on its distributional position and range relative to the first distribution. The choice of using the 13 notes of a one-octave range of the diatonic scale to represent the standard distribution was based on exploiting the "natural" perceptual structure of pitch. The octave provides both a familiar and a perceptually salient measurement unit for (1) assessing differences in location and range of the comparison stimulus distribution, and (2) describing "shape" in terms of where the longest strings of repeated notes fall *within* an octave.

The standard and comparison auditory histograms were each played as a 50-note string of 5.2 sec total duration, beginning with the lowest note category and ending with the highest. The histogram for the standard distribution was played first, followed by a .83-sec pause, followed by the histogram for the comparison distribution. This cycle was repeated twice, with a 1.67-sec pause be-

tween the pairs, following which the subject provided a dissimilarity rating for the pair of stimuli as described below.

Auditory quartile displays were a musical analogue of a box-and-whiskers drawing, which represented the minimum, the first, second, and third quartile, and the maximum scores of a distribution, as a five-note sequence of approximately 2.5 notes per second. These sequences were presented in pairs, with the display for the standard distribution presented first, followed by a .83-sec pause, followed by the comparison distribution. Two presentations of the pairs occurred on each trial, separated by a 1.67-sec pause, following which the subject made a dissimilarity rating. The pitch values representing the five notes were determined in the same manner as were the full auditory histograms, assigning middle C to the minimum value of the standard distribution, and an octave to its range. Subjectively, this produced a pair of arpeggio-like passages, in which the first passage always covered exactly one octave (e.g., C E G A C' for a relatively "normal" distribution), whereas the second could span more or less than an octave, depending on its range relative to the first distribution (e.g., C# E F# G D#', for a positively skewed distribution having a larger range than the standard). These types of displays were selected because they represent variants on a relatively familiar musical prototype (the arpeggio, albeit with an "extra" note), make use of the octave as a perceptual unit, and are sufficiently brief (in both time and number of notes) so that they can be easily encoded as a musical phrase. It seemed possible that this ease of encoding (which was supported by the observation that several subjects appeared to be covertly singing, whistling, or humming the passages as they performed dissimilarity ratings) might offset lack of detail about distribution shape attributable to the greater abstraction of data.

Visual histograms involved the visual display of a tick mark x-axis and a 12-category histogram in which frequency was coded by the length of vertical stacks of asterisk symbols. These histograms were displayed on the computer screen in text mode. On each trial, the histogram for the "standard" distribution was displayed for 3 sec, followed immediately by a 3-sec display of the "comparison" distribution histogram. Three cycles of alternation between the standard and the comparison histograms occurred (a total of 9 sec of viewing time), following which the screen was blanked and the subject made a dissimilarity rating.[1] Subjectively, the three cycles of the distributions produced an animation-like effect in which the distribution moved (with central tendency differences) and changed shape (with variability, skewness, and kurtosis changes) as displays were alternated, while the tick mark axis remained present throughout the histogram changes. Figure 1 shows examples of the visual histogram displays. The topmost display represents Stimulus 5, a relatively normal distribution with a midrange central tendency. The middle display represents Stimulus 3 (a relatively rectangular dis-

Table 1
Statistical Parameters of Distribution Stimuli

Stimulus	Shape	M	SD	Skewness	Kurtosis	Range
1	Normal	43.4	9.8	.07	−.51	42
2	Normal	44.6	9.7	−.07	.12	45
3	Flat	43.9	9.0	.29	−1.25	29
4	Skewed	43.8	7.7	1.06	.17	28
5	Normal	48.8	9.1	−.27	−.19	38
6	Normal	48.8	8.6	−.17	−.90	35
7	Flat	49.8	8.3	.16	−1.43	25
8	Skewed	50.0	9.5	1.12	.12	34
9	Normal	54.5	7.3	.48	.54	35
A	Normal	56.7	8.7	−.05	.13	40
B	Flat	54.8	8.8	.36	−1.20	28
C	Skewed	54.8	9.9	.82	−.69	34

Examples of Visual Histograms

```
                    *
                    *
                  * *
                *** *
Stimulus 5      *** *
"normal"        *** *
                **** *
             *  ****** *
             *  ********
                ************
!....!....!....!....!....!....!....!....!....!....!
```

```
Stimulus 3      *
"flat"          **
                **
                ***    *
                **** * *
                *******
                ********
                ********
                ********
!....!....!...,!....!....!....!....!....!....!....!
```

```
                    *
                    *
                    *
Stimulus C          *
"skewed"            *
                    *
                    *
                    *
                    *
                  * *
                  * **      *
                  ****       *
                  ***** * **
                  ***** ****
                  ***** *****
!....!....!....!....!....!....!....!....!....!....!
```

Figure 1. Examples of visual histogram displays.

tribution having a lower mean) scaled relative to Stimulus 5. The bottom display, also scaled relative to Stimulus 5, represents Stimulus C, which is positively skewed and has a higher mean.

Task

The subjects sat in front of a Gateway 2000 386SX computer, which was placed on a desk in a normally illuminated office, and which served as the stimulus presentation and data acquisition system. During each session, the subjects were presented with 132 comparison trials, within which each stimulus distribution was paired as the "standard" stimulus with the 11 other "comparison" stimuli, exactly once. Following the presentation of each stimulus pair (two alternations of the auditory stimuli or three alternations of the visual stimuli), the subjects used the computer keyboard to enter a whole number representing the judged dissimilarity of the stimulus pairs on a 1–10 scale, where 1 indicated *highly similar* and 10 indicated *highly dissimilar*. The subjects were instructed to base their judgments on differences in central tendency, variability, and shape of the distribution as if they were making informal comparisons between actual samples of data in a research setting.

During the first of the three sessions, each subject received three practice trial blocks of 20 trials each—one block for each presentation modality. These practice blocks were used to acquaint subjects with the task and to give them experience with the range of variability among the stimulus pairs and use of the rating categories. Apart from the presentation of these practice blocks in the first ses-

sion, each session included only one presentation modality for data collection. The order of modalities across the three sessions was block randomized between subjects. The 12×11 asymmetric matrix of dissimilarity ratings obtained in each session included two dissimilarity judgments for each pair of distributions (one for when each of the pair members was used as the standard). Since preliminary inspection of these asymmetric matrices did not reveal any systematic effect of which pair member was used as a standard (e.g., correlations between matrix halves exceeded .80 for all subjects, and there were no consistent discrepancies in judgment of any particular pair that depended on which member was the standard), the matrix halves were summed to produce a single measure of dissimilarity for each subject between each of the 66 pairs.

RESULTS

The dissimilarity ratings from 1 subject correlated negatively with those from the 11 other subjects, indicating a probable misunderstanding of instructions. The data from that subject were excluded from the analyses described below.

Correlations Between Dissimilarity Ratings and Stimulus Parameters

First, we wished to determine the extent to which dissimilarity judgments would be influenced by specific parameters of the numeric distributions, and whether sensitivity to differences in particular parameters would differ among the three display modes. We therefore computed, for each subject and display modality, the Pearson correlation coefficient between the 66 dissimilarity ratings and the pairwise differences in mean, standard deviation, skewness, kurtosis, and range. For an overall index of sensitivity to the *combination* of all five parameter differences, we also computed, for each subject and display modality, the multiple correlation coefficient (R) for prediction of dissimilarity judgments from the five parameter differences. Mean values of each of these correlations, averaged for the 11 subjects, are summarized in Figure 2.

The mean multiple R of .61 for the visual histograms, compared with .52 for both the auditory histograms and auditory quartile displays, suggests a slight overall "superiority" for visual presentation. Preplanned comparisons by t tests showed that the differences in multiple R for the visual histograms and the other two modes were each significant $[t(10) = 2.96$ and $3.65, p < .025$ and .01, for auditory histograms and quartile displays, respectively]. Three of the 11 subjects produced a higher multiple R for auditory than for visual histograms, and 2 of the 11 subjects produced a higher multiple R for auditory quartile displays than for visual histograms. More important, however, is that Figure 2 clearly shows that differences in central tendency (mean) produced a far greater influence on dissimilarity ratings for both types of auditory presentation $(r = .36$ and .40 for auditory histograms and quartile displays, but only .06 for visual histograms). Differences in shape information (skewness and kurtosis), on the other hand, were more highly correlated with dissimilarity judgments for the visual histograms than for either

Mean Correlations Between
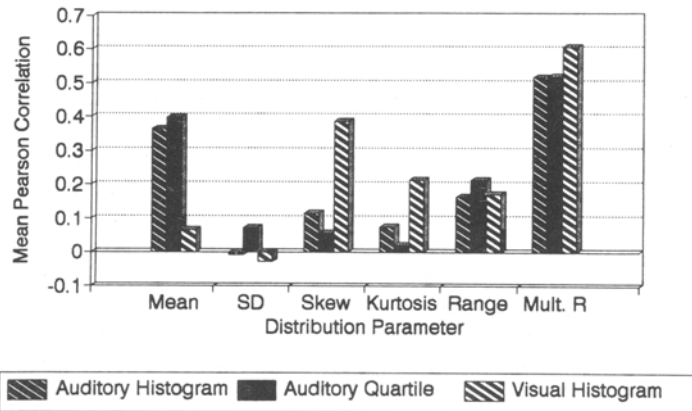### Judgments and Stimulus Parameters

Figure 2. Mean correlations between subjects' dissimilarity judgments and stimulus parameters.

of the auditory modes ($r = .39$ vs. $.11$ and $.06$ for skewness differences, and $r = .21$ vs. $.07$ and $.02$ for kurtosis differences). The essentially zero correlations between standard deviation differences and dissimilarity judgments seems likely to reflect the fact that little variation in standard deviations existed for the distributions used as stimuli.

The very low correlations between mean differences and dissimilarity ratings for the visual histograms were somewhat surprising, given that our subjects were relatively familiar with distribution displays of this type, and given our specific instructions to attend to differences in central tendency as well as shape. It should be noted that considerable individual differences existed among the correlations between mean differences and visual histogram dissimilarity judgments (maximum and minimum values were .63 and $-.28$). Since some subjects did perceive differences in central tendency nearly as well as they did with auditory presentation, whereas others seemed oblivious to positional information, it seems likely that extracting information about central tendency with the visual histograms may require a conscious attentional strategy that only a subset of our subjects employed.[2]

Perceptual Structure of the Averaged Dissimilarity Matrices

The means of correlations computed individually for each subject from data consisting of the sum of only two ratings for each stimulus pair may, due to the inherent noisiness of such data, provide a somewhat pessimistic view of the extent to which information about the stimulus parameters can be conveyed by each presentation mode. We thus performed a series of regression and multidimensional scaling analyses based on grand averages of the dissimilarity ratings across subjects. Although such a procedure may obscure potentially important individual differences, the use of an average dissimilarity matrix

for each display modality provides a model for an "average" subject. Such a model should reduce spurious effects from "noisy" individual subjects that could affect conclusions about differences in perceptual structure obtained from either regression models or MDS configurations.

Since the pattern of correlations from individual subjects suggested that central tendency dominated dissimilarity judgments with auditory presentation and that shape differences were more salient with visual presentation, it seemed likely that neither a single regression model with fixed coefficients nor a single MDS configuration would provide a very satisfactory fit to each of the three averaged dissimilarity matrices. However, to the extent that each of the three presentation modes conveyed information about the same set of stimulus properties, but with different degrees of *salience*, then a regression model containing the same *variables*, but different coefficients, might provide an adequate description of the three sets of dissimilarity measures. In addition, a common perceptual structure, differing only in the salience or weighting of each of the perceptual dimensions, might be extracted from an individual differences scaling (INDSCAL) procedure (Carroll & Chang, 1970). In the present case, such a perceptual configuration (or rotation of it) should ideally correspond to the statistically important properties along which the stimulus distributions varied (e.g., central tendency, skewness, and kurtosis).

Regression analysis. A standard forward regression procedure (SPSS-PC) was used, with the mean dissimilarity judgment (across subjects) for each stimulus pair as the dependent variable, and differences in mean, standard deviation, skewness, kurtosis, and range as potential predictors. Regression models that included the same predictor variables (differences in the mean, skewness, kurtosis, and range) were obtained for both auditory histograms and visual histograms. While the SPSS default cri-

Presentation Mode	M	Skewness	Kurtosis	Range
Auditory histograms	.63	.32	.16	.22
Auditory quartiles	.65	.19	.10	.31
Visual histograms	.21	.63	.40	.21

teria for variable inclusion produced a model for the auditory quartile display judgments that included only the mean differences and range difference, modeling those data with the same four variables (mean, skewness, kurtosis, and range) as those for the other modalities provided a reasonably acceptable fit, with each variable except kurtosis making a significant ($p < .05$) contribution to the prediction. The standardized beta coefficients for these regression analyses are summarized in Table 2.

It should be noted that the multiple R values that describe the ability of these four stimulus parameters to predict dissimilarity judgments were of nearly identical magnitude—.69 for auditory histograms, .70 for auditory quartile displays, and .73 for visual histograms. It is clear that differences in the mean had greater weighting for the auditory presentation modes, whereas shape differences were weighted more highly with the visual histograms. Nonetheless, information about central tendency, variability range, and shape made independent contributions to dissimilarity, regardless of the modality of presentation. The only qualification is reasonably attributable to the greater degree of abstraction in the five-note auditory box-whiskers—differences in kurtosis may have been obscured relative to the more "complete" auditory and visual histogram displays.

INDSCAL analysis. The three average dissimilarity matrices were submitted to INDSCAL analysis (SPSSX ALSCAL package, Version 3.0), treating each matrix as a separate "subject." Stress values (Kruskal's Stress Formula 1) for the derived configuration in three dimensions were .115 for the auditory histogram matrix, .135 for the auditory quartile displays, and .110 for the visual histograms. R^2 values (proportion of variance of the disparities accounted for by the distances in the configuration) were .871, .850, and .888 for the three matrix types. Collectively, this pattern suggests a reasonably good fit in three dimensions; it is slightly less good for the quartile displays than for either of the histograms. In Figure 3, the stimulus coordinates for the first two dimensions of the INDSCAL solution are plotted; in Figure 4, the coordinates for Dimension 3 are plotted against those for Dimension 2.

Inspection of these plots makes it clear that Dimension 1 represents perceived differences in central tendency. The four stimuli with the highest means (9, A, B, and C) are grouped on the left, the four stimuli with the midlevel means (5, 6, 7, and 8) occupy the center of the display, and the four stimuli with the lowest means (1, 2, 3, and 4) occupy the rightmost region. The Pearson correlation between the stimulus values for Dimension 1 and the stim-

ulus means for the 12 stimuli was −.96. Figures 3 and 4 show that distribution shape information is shared by those dimensions. In Figure 1, all the "normal" stimuli are located at the bottom of the plot, all having negative values on Dimension 2. The "skewed" stimuli (4, 8, C) are in a layer between .5 and 1; the stimuli at the top of Figure 1 (3, 7, B) are the "flat" stimuli. In Figure 4, the three shape categories form distinct clusters, with the skewed stimuli at the bottom right, the flat stimuli at the top right, and the normal stimuli at the middle left. Stimulus 6, which is somewhat of an outlier, is located closer to the "flat" stimuli than to the five other "normal" stimuli. However, Table 1 shows that this stimulus does indeed deserve placement in that location, since even though it was generated by the statistical routine that produced the other "normal" distributions, it has a kurtosis value of −.90, which is more platykurtic than any other stimuli, other than those generated by the routine that generated the "flat" distributions. Numerically, correlations between

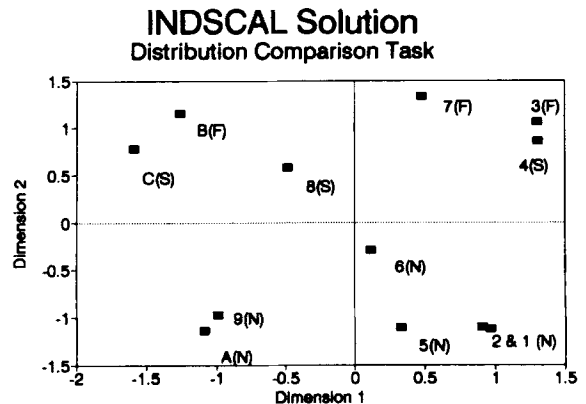**INDSCAL Solution**
Distribution Comparison Task

Figure 3. INDSCAL configuration obtained from the mean dissimilarity matrix (first two dimensions). Parentheses next to the stimulus label indicates the shape characteristic of the distribution: N, normal; F, flat; S, skewed.
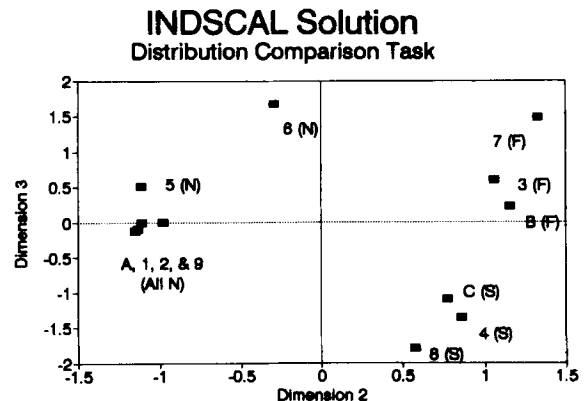
INDSCAL Solution
Distribution Comparison Task

Figure 4. INDSCAL configuration obtained from the mean dissimilarity matrix (Dimensions 2 and 3). Parentheses indicate the shape characteristic of the distribution: N, normal; F, flat; S, skewed.

(1) kurtosis and range and (2) Dimensions 2 and 3 confirm the sharing of shape characteristics among these dimensions. The Pearson correlation between Dimension 2 and Skewness was .57, and between Dimension 2 and kurtosis, it was −.62. Dimension 3 correlated −.79 with skewness, and −.58 with kurtosis. In summary, the perceptual structure that was produced by the INDSCAL routine appears to be interpretable, as well as indicative of the use of each of the important parameter variations that were built into our synthesized distributions.

Figures 5 and 6 present the "subject space" that depicts the relative weighting or salience of the INDSCAL dimensions among the three "subjects"—which in this case are not subjects but a *task space* for the averaged matrices across all subjects for each of the three display configurations.

The location of each of the tasks relative to the positive diagonal of these plots represents the relative "salience" of pairs of the MDS dimensions among the three tasks. In Figure 5, Dimension 1 (the central tendency dimension) is plotted against Dimension 2 (the first shape dimension), and in Figure 6, Dimension 1 is plotted

Task ("Subject") Space
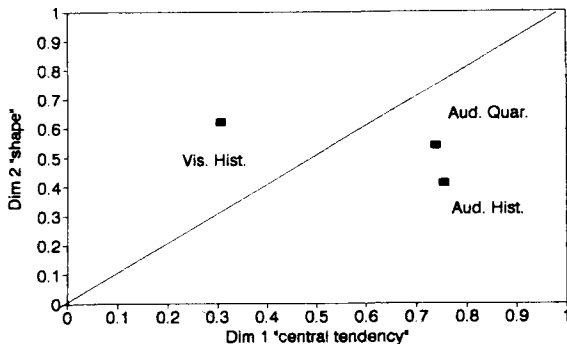### For Three Presentation Modes

Figure 5. Task space for the INDSCAL solution, Dimension 1 versus Dimension 2.

## Task ("Subject") Space
For Three Presentation Modes

Figure 6. Task space for the INDSCAL solution, Dimension 1 versus Dimension 3.

against Dimension 3 (the second shape dimension). For each plot, it can be seen that visual histograms are relatively dominated by shape (Dimensions 2 and 3), and the auditory displays by central tendency (Dimension 1), which is precisely the pattern anticipated from the results of our previous regression analyses.

MDSCAL analysis of the quartile display matrix. It was noted previously that one should expect some loss of detail of distribution shape in a highly abstracted display that only provides range and quartile information. Such degradation might account for the slightly less good fit to the common INDSCAL solution of the auditory quartile data as opposed to the other matrices, as well as the ability to model the dissimilarity judgments for the quartile displays by mean and range alone. This leaves the question of how well the auditory quartile displays allow discrimination among distributions of different shapes. To answer this, we performed a separate MDS analysis, using a traditional nonmetric MDS approach (again using the SPSSX ALSCAL package) for the auditory quartile data matrix alone. This time, an adequate fit occurred for a two-dimensional, rather than the three-dimensional, solution, with a stress value of .129 (R^2 = .887).[3] Figure 7 displays the obtained configuration.

Dimension 1 of the obtained configuration is clearly a central tendency dimension, with Stimuli 1, 2, 3, and 4 on the right, 5, 6, 7, and 8 in the center, and 9, A, B, and C on the left. Dimension 2 separates the normal stimuli from the others, with Stimuli 1, 2, 5, 6, 9, and A having positive values. The skewed and flat distributions all have negative values, and although there is a tendency for the most platykurtic stimuli to have lower values, there is no complete separation of these groups (specifically Stimuli 3 and 4 are in the "wrong" order on Dimension 2 for obtaining nonoverlapping clusters). Thus, these findings suggest that the five-note quartile displays did not provide quite as "rich" information about distribution shape, yet they did provide sufficient cues to allow subjects to discriminate between the relatively "normal" distributions and those that markedly departed from a Gaussian shape.
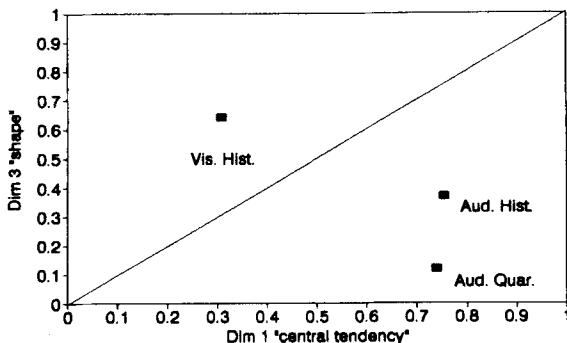
DISCUSSION

Our findings demonstrate that the auditory analogues of traditional visual graphic depictions of data are effective in conveying information about distribution central tendency, variability, and shape to observers who had never been previously exposed to, much less trained in the use of, auditory representations of data. Our data further showed that for the particular distributional stimuli and relatively crude display formats used in our study, central tendency and range information was conveyed more vividly by the auditory displays than by the more visual histograms, whereas distribution shape differences were more salient with visual depiction. It should be noted, however, that since we have made no formal attempt to optimize the structural properties of any of these

MDS Solution in 2 Dimensions
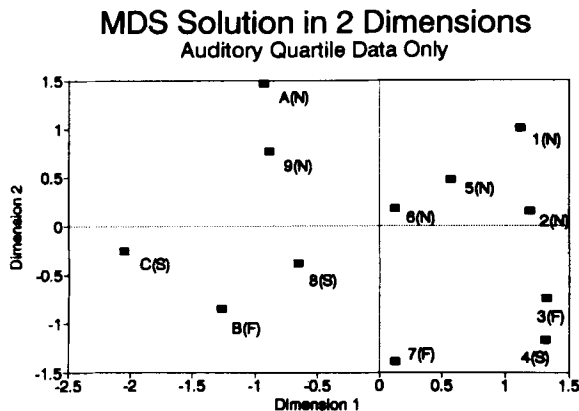### Auditory Quartile Data Only

Figure 7. Traditional MDS configuration obtained for the two-dimensional solution of the auditory quartile dissimilarity matrix only. Parentheses indicate the shape characteristic of the distribution: N, normal; F, flat; S, skewed.

display types for maximum efficiency of information transfer, it is not clear that these intermodal differences will prove to be completely generalizable phenomena. For example, relatively poor performance for judging central tendency differences among the visual histograms might reflect the particular way in which we displayed them. The sequential display, which subjectively provided a strong apparent motion cue about changes in distribution position to some subjects (and to the experimenters who designed the displays), apparently did not contribute to dissimilarity judgments for other subjects. The static, simultaneous visual display of the two distributions, as, for example, in the case of overlapping frequency polygons (line graphs, as opposed to histograms), might provide a more salient representation of the differences in central tendency that would be less likely to be unattended or discounted.

The eventual acceptance of nonvisual alternatives or supplements to visual graphic displays of data will depend on more than the potential of a novel display format to adequately convey quantitative information. One must consider the processing load imposed by the task of scanning such displays, of feeling them, or of listening to them—both in terms of the actual processing resources required and in terms of "perceived effort." Although we did not include a formal subjective "product evaluation" as part of this initial study, the informal comments of our subjects strongly suggested that the three display formats we used in this study differed substantially in effort required, with the auditory histogram judgments requiring the most effort and the lowest feelings of confidence about the ratings of dissimilarity. This may be attributable to the relatively large auditory memory load imposed by that task. The auditory quartile sequences (which were regarded as quite effortless to judge) fit well into the temporal span of auditory sensory memory, but the 5.7-sec sequence of 50 notes in the histogram displays may press those limits. Should some variant of auditory histograms eventually be considered as part of a specialized software pack-

age for the visually impaired, for example, careful consideration should be given to optimization of the temporal characteristics of the displays. Despite their greater level of abstraction and the resulting loss of detail of distribution shape properties, we believe that for exploratory data analysis, the quartile or box-whisker approach may have greater utility (and receive greater acceptance among users) than full histogram displays do, when used either as enhancements for visual data displays or as a "quick and dirty" procedure for obtaining an overview of salient distributional properties. For example, we are currently investigating the utility of slightly modified auditory "box-whisker" displays, in which central tendency information (the median) is represented as an initial sustained note of a six-note melody, and the quartiles or "hinges and whiskers" are defined by a more rapid and staccato five-note "arpeggio" that follows the sustained median. Temporal and rhythmic modification such as these may be useful in "tuning" the relative salience of different properties of data to best fit their relative importance in the task for which they are being used.

REFERENCES

BLATTNER, M., GREENBERG, R., & KAMEGAI, M. (1990). Listening to turbulence: An example of scientific audiolization. In *Proceedings of the ACM SIGCHI 90 Workshop on Multimedia and Multimodal Interface Design* (pp. 1-8). New York: Association for Computing Machinery.

BLY, S. (1982). Presenting information in sound. In *Proceedings of the CHI '82 Conference on Human Factors in Computer Systems* (pp. 371-375). New York: Association for Computing Machinery.

CARROLL, J. D., & CHANG, J. J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition. *Psychometrika*, 35, 238-319.

GARBIN, C. P. (1988). Visual-haptic perceptual non-equivalence for shape information and its impact upon cross-modal performance. *Journal of Experimental Psychology: Human Perception & Performance*, 14, 547-553.

SCALETTI, C., & CRAIG, A. (1991, February). *Using sound to extract meaning from complex data.* Paper presented at the Society of Photo-optical Instrumentation Engineers Conference 1259, San Jose, CA.

NOTES

1. The choice of allowing three "looks" at these durations (as opposed to two "listenings" for the auditory stimuli) was made on the basis of subjective impressions of the experimenters and some volunteer pilot subjects. This temporal patterning of the visual stimuli provided a sufficient opportunity to sample the stimulus displays prior to rating their dissimilarity, which seemed to be psychologically equivalent to the stimulus sampling time for the tone displays.

2. It should be noted that the first author, who served as a pilot subject (but whose data are not included due to unrepresentative familiarity with the stimulus distributions that he had personally synthesized and selected), provided dissimilarity ratings that were much more strongly correlated with the mean for visual presentation than for either auditory mode ($r = .72$ vs. .63 and .62). It is possible that familiarity with the range of variation along specific stimulus attributes may be more critical for visual graphics than for auditory displays.

3. We performed individual matrix analyses for all three tasks, but since the individual three-dimensional solutions for both auditory histograms and visual histograms could each be rotated to solutions that did not substantially differ from the INDSCAL solution, we chose not to include them here. What is important to note, however, is that neither the auditory histograms nor the visual histograms could be scaled with acceptable stress values in *less* than three dimensions.