

# An investigation of data collection methods for auditory stimuli: Paired comparisons versus a computer sorting task

TERRI L. BONEBRIGHT

*University of Nebraska, Lincoln, Nebraska*

Two experiments were performed to determine whether a computer sorting task could be used to collect data for auditory stimuli. Within each experiment, subjects were assigned to either a paired comparison task or a computer sorting task. For the first experiment, the stimuli were 18 graph representations constructed to vary on three parameters (shape, frequency, and amplitude). The stimuli for the second experiment were 30 sentences produced by actors portraying five emotions (anger, fear, happiness, sadness, and neutral). Multidimensional scaling solutions were produced and comparisons were made between the paired comparison and computer sorting task in each experiment. The results suggest that the computer sorting task is a viable alternative to the paired comparison task traditionally used for the investigation of the perception of auditory stimuli.

Sorting tasks are routinely used to collect data in a variety of fields within psychology. This type of task has been assumed to be more efficient in terms of time spent collecting data and more motivating for subjects than other tasks. To test this assumption, researchers have performed comparison studies to determine whether sorting tasks provide data that are as accurate as or better than data obtained with other methods. Their results have provided support for the validity of sorting tasks as a research method within such diverse areas as measurement of cognitive content (Hirschman & Wallendorf, 1982), measurement of feeling states (Newton, Prensky, & Schuessler, 1982), career guidance instruments (Jones, 1983), and Stroop test presentation (Taylor & Clive, 1983).

Within perceptual research, visual and tactile stimuli have also been examined by using sorting techniques to obtain data appropriate for multidimensional scaling (MDS) solutions used to explore the perceptual dimensions of such stimuli (Garbin, 1988; Garbin & Bernstein, 1984; Hollins, Faldowski, Rao, & Young, 1993). However, auditory stimuli are traditionally examined by using paired comparison tasks rather than sorting tasks, because of the temporal rather than spacial nature of the stimuli (Flowers & Hauer, 1995). Thus, if a researcher is interested in investigating a large number of auditory stimuli, the number of subjects necessary and the time commitment of those subjects can prove to be restrictive when the only option for data collection is a paired comparison task. In such circumstances, sorting tasks could provide an alternative method if subjects could actually perform the sorting of auditory stimuli. Therefore, the purpose of the present research was to determine whether a computer sorting task

could be used successfully to investigate the perception of auditory stimuli.

Two experiments, each with a different set of auditory stimuli, were performed to compare a paired comparison task and a computer sorting task. In Experiment 1, the stimuli were auditory graphs, computer produced to vary on three known parameters (shape, frequency, and amplitude). In Experiment 2, the stimuli were actors' portrayals of vocal affect that varied for the type of emotion expressed (happiness, anger, sadness, fear, and neutral). Through comparison of the two tasks for both types of auditory stimuli, evidence could be gained for assessing the validity of using a computer sorting task.

## EXPERIMENT 1

### Method

**Participants.** Twenty-six introductory psychology students participated (15 women, 11 men; mean age, 23.4) as an alternative means of obtaining course credit. Half of the participants performed the paired comparison task; the other half performed the computer sorting task.

**Materials.** Eighteen auditory graph representations were generated by computer, using 100 numbers each to produce wave-form samples. Wave forms varied in shape (sine, square, or a composite of the two), frequency (4, 6, or 8 cycles per 100 points, called low, medium, and high, respectively), and amplitude (low or high). These data sets were then used to generate auditory graphs through a computer speaker as strings of 100 musical notes with a total duration of 6 sec for each graph. Data values were mapped onto a diatonic musical scale so that the largest peak-to-peak amplitude changes in the wave forms in the stimulus set were represented auditorily by a two-octave shift. This mapping procedure was similar to that used by Flowers and Hauer (1992, 1993, 1995) for similar types of auditory data representations.

For the paired comparison task, the 18 stimuli were randomized to form two lists which were counterbalanced across subjects, and a Gateway 2000 386X was used for stimulus presentation. For the sorting task, a Macintosh Quadra with the Canary Bioacoustics Re-

Correspondence should be addressed to the author at the Department of Psychology, DePauw University, Greencastle, IN 46135 (e-mail: tbone@depauw.edu).

search Program (1982; Version 1.1) was used for digitization and presentation of stimuli.

**Design and Procedure.** Participants were randomly assigned to either the paired comparison task or the computer sorting task. For both data collection conditions, participants were presented with all 18 auditory graphs to familiarize them with the range of variation among the stimuli.

In the paired comparison task condition, groups of 3–5 students participated in 2-h sessions which included a 10-min break. Participants listened to each pair of graphs twice and then made a similarity rating using a 1 (*not similar*) to 9 (*very similar*) Likert scale. A practice task with auditory graphs not included in the actual stimulus set was performed prior to the actual data collection in order to answer any questions participants had about the procedure.

In the sorting task condition, individual students listened to the stimuli on headphones during the hour procedure. Before the auditory graph sorting task began, a practice sorting task, consisting of a set of animal sounds, was used to teach the participants how to move the icons on the screen and how to open and play the auditory files. Once the practice task had been completed successfully, the participants were presented with a screen containing all icons representing the auditory graph files, which were placed in two columns on the left side. They were told to begin with the icon at the bottom of the left column, listen to it, and then move it to the right side of the screen. The participants then listened to the next stimulus from the bottom of the column and moved it into the same column as that with the first stimulus if they judged it to be similar or moved it to a new column if they judged it to be different from the first stimulus. This procedure was continued until all stimuli were placed in groups. The participants were allowed to listen to any of the stimuli as often as they liked and to change the group members to make sure that the members of each group reflected the similarities and differences among the graphs. They were also told that there was no right or wrong way to sort the stimuli and that they could use whatever rules for group members they deemed appropriate, except that a “group” had to consist of a minimum of two graphs. Once the sorting was complete, participants listened to all the stimuli again as a final check for group membership and made any adjustments that they felt were necessary. The final step in the procedure required the participants to answer questions about the strategy they used to sort the stimuli.

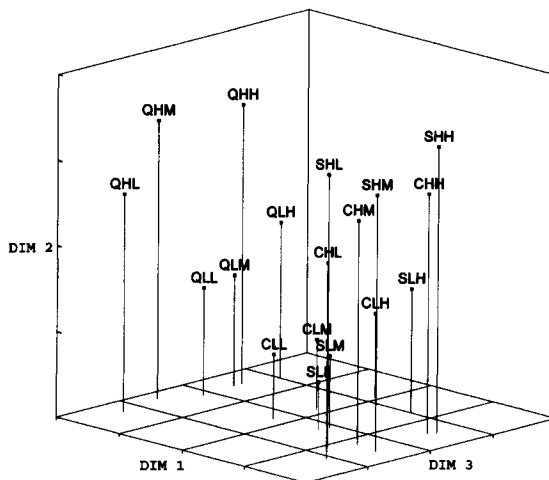
## Results and Discussion

MDS solutions were generated for both the paired comparison data and the sorting data. For the paired comparison data, the mean rating for each pair was used to produce the lower off-diagonal  $18 \times 18$  composite dissimilarity matrix. For the sorting data, individual dissimilarity matrices were compiled by using a 0 to indicate that the subject had sorted the pair into the same group and a 1 to indicate that the subject had placed the pair into different groups (Shiffman, Reynolds, & Young, 1981). The individual matrices were then aggregated to produce a composite dissimilarity matrix. MDS solutions of the two composite matrices with 1–6 dimensions were obtained with ALSCAL (Young & Lewyckij, 1979). The  $R^2$  and stress values for Dimensions 1–6 for the sorting data suggested that a 3-dimensional solution produced the best fit with the data ( $R^2 = .75, .81, .93, .98, .99, .99$ ; stress =  $.29, .19, .10, .04, .02, .01$ ). However, the  $R^2$  and stress values for the paired comparison task suggested that a 2-dimensional solution produced the best fit ( $R^2 = .77, .91, .96, .96, .97, .98$ ; stress values =  $.27, .14, .08, .06, .05, .04$ ).

To compare the solutions from the paired comparison and sorting tasks, the stability of the 3rd dimension for the paired comparison data was considered. To check for stability, split-half samples of the matrix were formed comprising 6 and 7 subjects. These matrices were scaled, using ALSCAL, and the  $R^2$  values (.93) and the stress values (.09) were found to be the same for the 3-dimensional solutions, which suggests a high degree of stability. As an additional check, these two solutions and the full matrix were orthogonally rotated to positions of maximum congruence (Harman, 1976). The rotation of the three solutions revealed coefficients of congruence (CCs), which can be interpreted like a correlation coefficient (Garbin, 1990), of above .9 for all comparisons, providing further support for the similarity of the dimension by dimension positioning of the stimuli within the 3-dimensional space for the split-half solutions and the full matrix.

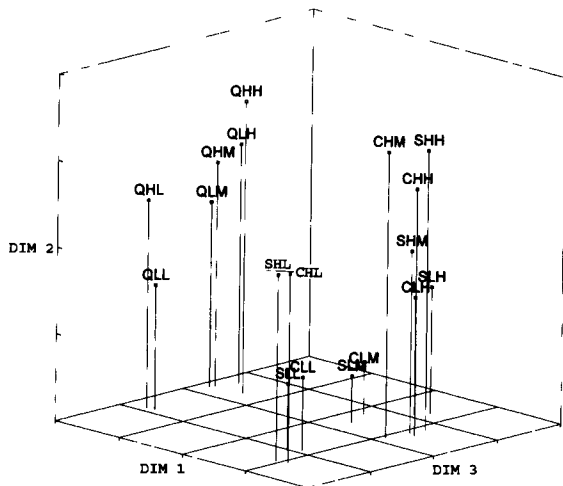
With the evidence for stability provided by the previous analyses, the comparison of the paired comparison solution with the sorting task solution was considered for 3 dimensions. These two solutions were rotated to positions of maximum congruence, and the CCs were used to make dimension by dimension comparisons. The CCs were .96, .86, and .78 for Dimensions 1–3, respectively, which suggests good stability for the first 2 dimensions and moderate stability for the 3rd.

Inspection of Figures 1 and 2 reveals that an interpretation of the positioning of the stimuli within the stimulus space for both data sets is consistent for all 3 dimensions. When one views Dimension 1 from left to right, it can be seen that this dimension separates the square-wave forms from the sine and composite wave forms. Dimension 2, considered from the bottom to the top of the cube, reveals positioning of the stimuli from low to high amplitude. Finally, Dimension 3, viewed from the front to the back of the cube, shows the stimuli arranged according



Key for Stimulus Labels: The letters for the three letter labels in order represent type (Q = square, S = sine, & C = composite), amplitude (L = low & H = high), and frequency (L = low, M = medium, & H = high).

Figure 1. Paired comparison task for auditory graph representations.



Key for Stimulus Labels: The letters for the three letter labels in order represent type (Q = square, S = sine, & C = composite), amplitude (L = low & H = high), and frequency (L = low, M = medium, & H = high).

Figure 2. Computer sorting task for auditory graph representations.

to low, medium, and high frequencies. The differences in the two solutions appear to lie primarily in the closeness of the groupings of the stimuli. For the sorting task, the stimuli tend to be clustered into regions while the paired comparison task produced a more evenly distributed arrangement. However, these differences do not detract from the overall similarity in the stimulus spaces for the two tasks.

## EXPERIMENT 2

### Method

**Participants.** Participants were 155 introductory psychology students who received course credit for performing the experimental task. All participants were native English speakers, ranging in age from 18 to 42 years old (mean, 20.5). Data for the paired comparison task were collected from 113 participants (44 males and 69 females), with data for 19 participants removed because of missing values. Data for the sorting task were collected from 42 participants (21 males and 21 females), with data for 6 participants discarded because of missing values.

**Materials.** Vocal affect stimuli were produced by three female and three male English speaking actors who performed readings of a story that contained emotionally neutral content while portraying anger, happiness, sadness, fear, or neutral emotion. A single sentence was extracted for digitization and presentation. This sentence (*Everyone tried to convince him that the home repairs could wait until after the vacation*) was chosen because of its position within the passage and its general intonation pattern. It was also believed that the verbal content of this particular passage worked well for the portrayal of the five emotional states.

For the paired comparison task, a PAL (Personal Acoustics Lab) system with a Gateway 200 386x was used for digitization and playback purposes. All 30 stimuli were used for the paired comparison task, resulting in 900 pairs for all possible pairs (each stimulus paired with every other stimulus in both presentation orders, and each stimulus paired with itself). One half of this set of stimuli (one presentation order) was then divided into 5 sets of 90 pairs by using a modified random order with the following constraints: if an actor was paired with him-/herself, that actor was not allowed

in the next pair; if two males or females were in a pair, the next pair could not have the same combination; and if two stimuli with the same emotion were in a pair, the next pair could not have that emotion. This set of stimuli was then reversed in presentation order for each pair as well as for the entire list, resulting in 10 lists of 90 pairs each. Finally, a single practice set of 15 pairs was added to the beginning of each list, and an additional 15 pairs were duplicated within each list to serve as reliability checks. Thus, each of the 10 lists had a total of 120 stimulus pairs, and groups of participants were randomly assigned to a single list for the experimental procedure.

For the sorting task, as in Experiment 1, the equipment used for digitization and presentation of the stimuli was a Macintosh Quadra with the Canary Bioacoustics Research Program (Version 1.1).

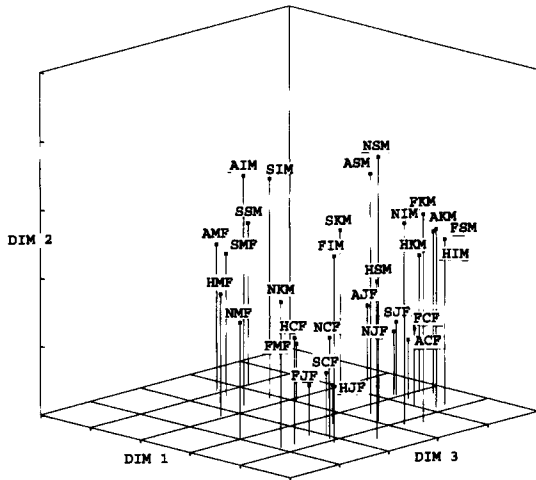
**Design and Procedure.** The experimental sessions for the paired comparison task took place over 5 weeks with groups of 4–8 students participating at each session. The participants were told that they would be listening to pairs of vocal stimuli which were all made from the same sentence. After they had listened to both sentences of each pair, they were to rate them on a 7-point Likert scale ranging from 1 (*not at all similar*) to 7 (*extremely similar*). The instructions also informed the participants that they would have a chance to practice the task on the first 15 pairs of sentences and to ask questions after the completion of the practice trials. After the participants had completed rating the 120 pairs, they were asked to list the types of cues that they had used to make their decisions.

The sorting task for this experiment followed the same basic procedure as that in Experiment 1. Students participated individually and performed a practice task with animal sounds to become familiarized with the computer and software. They were then presented with all the sentences, to acquaint them with the range of the stimuli. The procedure for the sorting task for the sentences allowed the participants to replay and move any icon at any time in order to make sure that the decisions they made about group membership reflected the relationships among stimuli. No constraints were placed on the number of groups, but participants were constrained to forming groups consisting of at least two stimuli. As in Experiment 1, the participants were encouraged to use whatever strategy they felt was appropriate to make the group membership judgments. Once the sentences were sorted, the participants were required to listen to all the sentences in each group as a final check to make sure that they were satisfied with their groupings. The experimenter then recorded the numbers of the stimuli and asked the participants to explain the basis that they had used for sorting the sentences.

### Results and Discussion

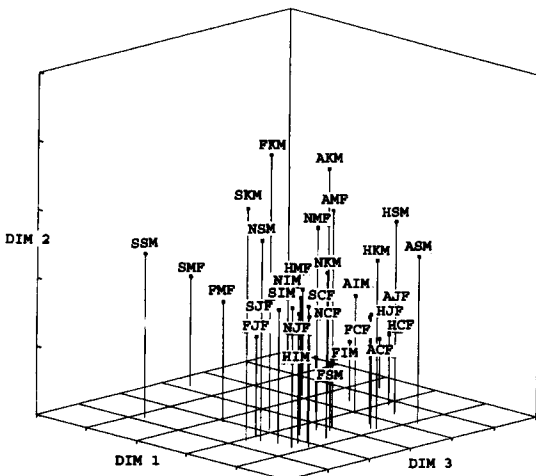
MDS solutions were produced for the paired comparison and the sorting data according to the same procedure as in Experiment 1. The  $R^2$  and stress values for Dimensions 1–6 for both solutions suggested that a 3-dimensional solution produced the best fit (for paired comparison data,  $R^2 = .32, .67, .81, .88, .93, .95$ , and stress =  $.50, .24, .15, .10, .07, .06$ ; for sorting data,  $R^2 = .72, .81, .86, .88, .90, .92$ , and stress =  $.31, .20, .15, .13, .10, .08$ ). The two 3-dimensional solutions were orthonormally rotated to positions of maximal congruence, and the obtained CCs revealed that the two solutions were moderately similar on the first dimension (Dimension 1 CC =  $.60$ ) but were not similar on the other 2 dimensions (Dimension 2 CC =  $.40$ ; Dimension 3 CC =  $.02$ ).

Inspection of Figures 3 and 4 reveals a positioning of stimuli that favors an interpretation dependent on the clusters or neighborhoods found in the 3-dimensional spaces, rather than descriptions of stimuli as they fall on each dimension. For the paired comparison data, the stimuli along Dimension 2 fall into regions of male and fe-



Key for Stimulus Labels: The letters for the three letter labels in order represent emotion (H = happy, S = sad, F = fear, A = anger, & N = neutral), actor (K, C, S, J, I, & M) and gender (M = male & F = female).

Figure 3. Paired comparison task for vocal affect stimuli.



Key for Stimulus Labels: The letters for the three letter labels in order represent emotion (H = happy, S = sad, F = fear, A = anger, & N = neutral), actor (K, C, S, J, I, & M) and gender (M = male & F = female).

Figure 4. Computer sorting task for vocal affect stimuli.

male actors which are roughly clustered into groups of portrayals by individual actors. There is no evidence that the positions of the stimuli correspond to the emotions expressed in the portrayals within this solution. For the sorting task data, the stimuli are grouped into regions composed of emotions along Dimension 3 with the active emotions (happy and anger) positioned together to the right of the figure and the inactive emotions (sadness, fear, and neutral) positioned together at the left. Within each of these emotion regions, the stimuli tend to cluster into groups of male and female actors.

**GENERAL DISCUSSION**

The results from both studies suggest that an auditory sorting task can be performed by subjects using a com-

puter software package. In the case of the auditory graph representations, both solution spaces reproduce the parameters used to construct the stimuli. This provides strong evidence for the usefulness of both types of task for this set of auditory stimuli. However, for the vocal affect stimuli, the evidence to support the sorting task is not as clear. It does appear that the data from the sorting task reflect the emotional cues in the stimuli, whereas the paired comparison data does not seem to do so. The divergence of the solutions for these two tasks may be evidence for the perceptual ambiguity of the stimuli rather than a reflection on the ability of subjects to perform the computer sorting task for the auditory stimuli. Future research should concentrate on replication of such findings using different auditory stimuli and should also seek to determine the best software for this type of task.

**REFERENCES**

Canary Bioacoustics Research Program 1.1 [Computer software]. (1992). Ithaca, New York: Cornell Laboratory of Ornithology.

CLIFF, N. (1966). Orthogonal rotation to congruence. *Psychometrika*, **31**, 33-42.

FLOWERS, J. H., & HAUER, T. A. (1992). The ear's versus the eye's potential to assess characteristics of numeric data: Are we too visuo-centric? *Behavior Research Methods, Instruments, & Computers*, **24**, 258-264.

FLOWERS, J. H., & HAUER, T. A. (1993). "Sound" alternatives to visual graphics for exploratory data analysis. *Behavior Research Methods, Instruments, & Computers*, **25**, 242-249.

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

GARBIN, C. P. (1988). Visual-haptic perceptual nonequivalence for shape information and its impact upon cross-modal performance. *Journal of Experimental Psychology*, **14**, 547-553.

GARBIN, C. P. (1990). Visual-touch perceptual equivalence for shape information in children and adults. *Perception & Psychophysics*, **48**, 271-279.

GARBIN, C. P., & BERNSTEIN, I. H. (1984). Visual and haptic perception of three-dimensional solid forms. *Perception & Psychophysics*, **36**, 104-110.

HARMAN, H. H. (1976). *Modern factor analysis* (3rd ed.). Chicago: University of Chicago Press.

HIRSCHMAN, E. C., & WALLENDORF, M. R. (1982). Free-response and card-sort techniques for assessing cognitive content: Two studies concerning their stability, validity, and utility. *Perceptual & Motor Skills*, **54**, 1095-1110.

HOLLINS, M., FALDOWSKI, R., RAO, S., & YOUNG, F. (1993). Perceptual dimensions of tactile surface texture: A multidimensional scaling analysis. *Perception & Psychophysics*, **54**, 697-705.

JONES, L. K. (1983). A comparison of two self-directed career guidance instruments: Occu-sort and Self-directed search. *School Counselor*, **30**, 204-211.

NEWTON, R. R., PRENSKY, D., & SCHUESSLER, K. (1982). Form effect in the measurement of feeling states. *Social Science Research*, **11**, 304-317.

SHIFFMAN, S. S., REYNOLDS, M. L., & YOUNG, F. W. (1981). *Handbook of multidimensional scaling*. New York: Academic Press.

TAYLOR, A., & CLIVE, P. B. (1983). Two forms of the Stroop test. *Perceptual & Motor Skills*, **57**, 879-882.

YOUNG, F. W., & LEWYCKYJ, R. (1979). *ALSCAL-4: Users guide*. (Available from F. W. Young, Department of Psychology, University of North Carolina, Chapel Hill, NC 27514)