

1996

The effects of task demands on the equivalence of visual and auditory representations of periodic numerical data

Kimberly D. Turnage
University of Nebraska-Lincoln

Terri L. Bonebright
Depauw University

Dion C. Buhman
University of Nebraska-Lincoln

John H. Flowers
University of Nebraska-Lincoln, jflowers1@unl.edu

Follow this and additional works at: <http://digitalcommons.unl.edu/psychfacpub>

 Part of the [Psychiatry and Psychology Commons](#)

Turnage, Kimberly D.; Bonebright, Terri L.; Buhman, Dion C.; and Flowers, John H., "The effects of task demands on the equivalence of visual and auditory representations of periodic numerical data" (1996). *Faculty Publications, Department of Psychology*. 445.
<http://digitalcommons.unl.edu/psychfacpub/445>

This Article is brought to you for free and open access by the Psychology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, Department of Psychology by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

The effects of task demands on the equivalence of visual and auditory representations of periodic numerical data

KIMBERLY D. TURNAGE

University of Nebraska, Lincoln, Nebraska

TERRI L. BONEBRIGHT

DePaul University, Greencastle, Indiana

and

DION C. BUHMAN and JOHN H. FLOWERS

University of Nebraska, Lincoln, Nebraska

The equivalence of visual and auditory graphical displays was examined in two experiments. In Experiment 1, multidimensional scaling techniques were applied to paired comparison similarity judgments of both auditory and visual displays of simple periodic wave forms. In Experiment 2, a subset of perceptually similar pairs of wave forms was selected as the stimulus set for an AB-X discrimination task in both auditory and visual modalities. Results suggest much greater apparent visual-auditory equivalence for the similarity rating task than for the more difficult discrimination task, implying that one should consider the demands of the task when deciding whether auditory graphic displays are suitable alternatives to more traditional visual displays.

Advances in computer technology offer possibilities for data summary and presentation in highly nontraditional formats. Flowers and Hauer (1992, 1993, 1995) have demonstrated that the perceptual structures of vision and audition are highly similar for the representation of graphs depicting statistical distribution and time series functions. The present investigation was designed to extend previous research to a new form of data: periodic wave forms.

This project included both an assessment of visual-auditory perceptual equivalence for these graphs (based on similarity ratings) and an assessment of the equivalence of accuracy performance for the two modalities on a more demanding AB-X discrimination task. If auditory displays are to be used as substitutes for visual graphics, it is necessary to assess the degree of accuracy achievable by the two modalities for demanding tasks such as discriminating between different data sets. Knowing the relative performance levels of the two modalities for such discriminations is useful in decisions about the appropriateness of auditory graphics in general, the optimal design of auditory displays, and the appropriate measures for training people to use them.

EXPERIMENT 1

In this study, we assessed whether the degree of visual-auditory perceptual equivalence of these wave-form stim-

uli would parallel the perceptual equivalence reported for other types of graphs. If wave-form data are as suitable for auditory presentation as distributional and trend data are, visual and auditory perceptions of these stimuli should be very equivalent.

Method

Subjects. Twenty-six undergraduate psychology students who reported having normal or corrected-to-normal vision and hearing participated to fulfill a course research requirement. Thirteen subjects (balanced for gender) participated in each modality condition.

Materials. Eighteen data samples, each consisting of 100 numbers that described a periodic wave-form function, were computer generated. These 18 wave forms varied in shape (sine, square, and a composite, which was a weighted sum of the two), frequency (low, medium, and high, which were 4, 6, and 8 cycles per 100 points, respectively) and amplitude (low and high). The chart function of Microsoft Excel was used to produce visual graphs of these wave forms. (See Figure 1 for examples of visual graphs.) These visual graphs were transferred onto overhead transparencies and were presented with an overhead projector. The same wave forms were presented as auditory graphs through the speaker of an IBM-compatible 386 computer. The graphs were played 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 other types of data representations.

Procedure. The task of subjects in both the visual and the auditory conditions was to provide similarity ratings for all independent pairs of the 18 wave forms (153 pairs) using a 9-point scale (1 = *very different*; 9 = *very similar*). Before beginning the rating task, subjects were presented with all 18 graphs in a random order to familiarize them with the entire range of stimuli. Each subject

Correspondence should be addressed to K. D. Turnage, Department of Psychology, 209 Burnett Hall, University of Nebraska, Lincoln, NE 68588-0308.

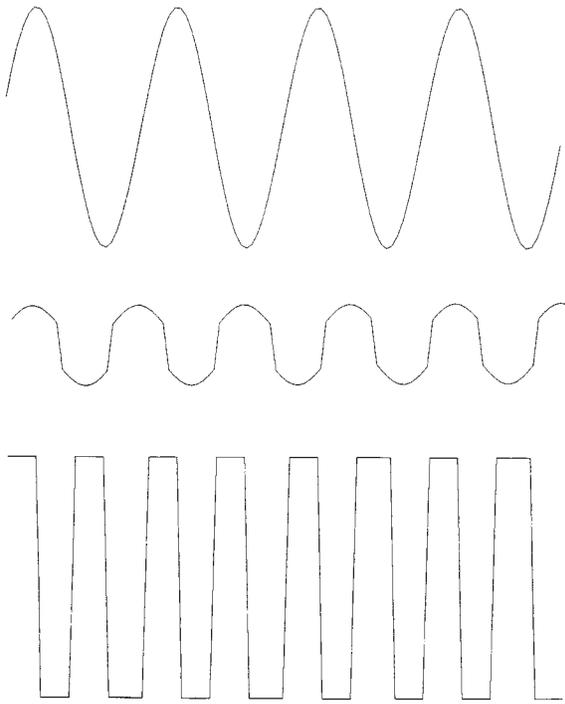


Figure 1. Examples of visual wave forms used in Experiments 1 and 2. **Top:** high-amplitude, low-frequency sine wave. **Middle:** low-amplitude, medium-frequency composite wave. **Bottom:** high-amplitude, high-frequency square wave.

then completed three practice trials with pairs of wave forms which were not among the 18 used as stimuli, in order to become familiar with the similarity rating process. For the rating task, each graph was presented twice, and the subjects were instructed to make their similarity judgments only after seeing or hearing each pair of wave forms twice. Each subject completed 153 test trials and received a 5-min break after rating one half of the stimulus pairs.

The visual graphs were presented with an overhead projector. Each graph was presented for 4 sec, with a 2-sec interval between graphs within a pair and a 4-sec interval between the first and second presentations of each pair. The auditory graphs were presented on the speaker of an IBM-compatible 386 computer as square-wave musical notes generated by the Microsoft QuickBasic Version 4.5 "play" function. The duration of each auditory graph was 6 sec, with a 2-sec interval between members of a pair and a 4-sec interval between the first and second presentation of each pair.

Results and Discussion

For each modality condition, a mean dissimilarity rating was obtained for each pair of graphs, and these dissimilarities were collected into the lower off-diagonals of two 18×18 composite dissimilarity matrices, one for visual subjects and one for auditory subjects.

Each of these composite dissimilarity matrices was submitted to multidimensional scaling (MDS) analysis. Analyses revealed that 3-dimensional solutions could be used to describe the perceived similarities and differences among the graphs for both visual ($R^2 = .857$, stress = .128) and auditory ($R^2 = .933$, stress = .099) subjects. (See Table 1 for R^2 and stress values in 1–4 dimensions for each modality.¹)

For the purpose of comparison, the two solutions were orthonormally rotated to positions of maximal congruence. After rotation, the visual and auditory solutions were compared dimension by dimension, using the coefficient of congruence (CC), which is interpreted like a correlation coefficient (see Cliff, 1966; Garbin, 1988, 1990). Coefficients of congruence revealed that the visual and auditory solutions were very similar on all three dimensions (Dimension 1 CC = .963; Dimension 2 CC = .978; Dimension 3 CC = .941).

Figures 2 and 3 illustrate the 1st and 2nd and 1st and 3rd dimensions, respectively, for the rotated auditory and visual MDS solutions. Examination of Figure 2 reveals that for both modalities the stimuli were perceived as four distinct neighborhoods. Clockwise from the upper right quadrants, the first neighborhood was defined by sine and composite high-amplitude wave forms; the second consisted of square high-amplitude wave forms; the third consisted of square low-amplitude waves; and the fourth was defined by sine and composite low-amplitude wave forms. Thus, Dimensions 1 and 2 for both modalities were interpreted as separating stimuli on the basis of shape and amplitude, respectively.

Figure 3 depicts the Dimension 1 \times Dimension 3 space for these wave forms for the two modalities. In this space, stimuli did not form neighborhoods. However, examination of Figure 3 reveals that for both the auditory and the visual subjects, high-frequency stimuli generally had positive values along the 3rd dimension, whereas low-frequency stimuli had negative values, and medium-frequency stimuli fell between the two.

The perceptual structures for vision and audition, based on similarity ratings, appeared very equivalent both in their interpretation and on the basis of the CC values. Although the overall interpretations for the visual and auditory perceptual structures were very similar, some within-neighborhood differences warrant mention. Within the first and fourth neighborhoods, which consisted of sine and composite high- and low-amplitude wave forms, respectively, there were greater distances among the stimuli in the auditory than in the visual perceptual space. Examination of Figure 2 suggests that this separation may have been based on greater perceived

Table 1
 R^2 and Stress Values of MDS Solutions for Vision and Audition in 1–4 Dimensions

No. of Dimensions	R^2	Stress
Vision		
1	.58	.45
2	.76	.22
3	.86	.12
4	.93	.07
Audition		
1	.76	.27
2	.91	.14
3	.95	.08
4	.95	.06

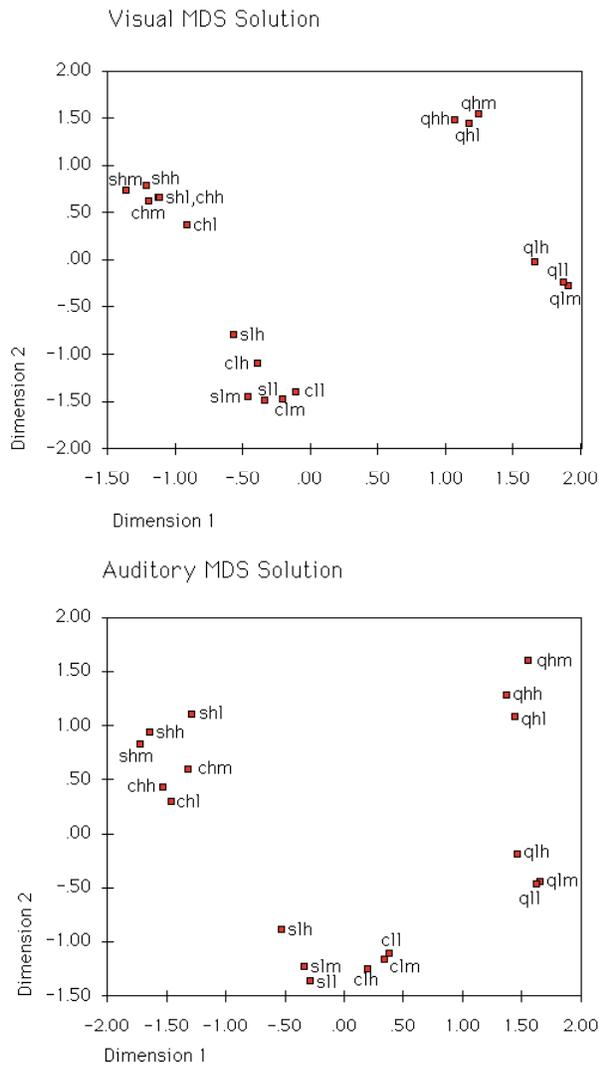


Figure 2. Visual and auditory MDS solutions for Dimensions 1 and 2. Stimuli are denoted according to shape (s, sine; q, square; c, composite), amplitude (h, high; l, low), and frequency (h, high; m, medium; l, low), respectively.

shape differences for audition than for vision. Note that the sine waves are all closer to one another than to the composite waves and vice versa in the auditory space, but not in the visual space. On the basis of these within-neighborhood differences, it would appear that very similar stimuli (i.e., the composite and sine waves) may have been more separable by audition than by vision.

These results suggest that on a task involving similarity ratings of pairs of stimuli, visual and auditory representations of the same stimuli are very equivalent. Although there were slight differences in perceived similarity within neighborhoods (particularly those defined by non-square waves) between the modalities, the between-neighborhood differences for the two modalities were very congruent; this congruence was also reflected by the high CC values for all three dimensions. These re-

sults, then, converge with results from previous experiments (e.g., Flowers & Hauer, 1992, 1993, 1995) to suggest substantial perceptual equivalence between vision and audition for numeric or graphical information.

EXPERIMENT 2

Results from Experiment 1 suggest substantial visual–auditory perceptual equivalence for these wave-form data. However, a distinction has been made in the perceptual literature between structural perceptual equivalence and functional performance equivalence (e.g., Garbin, 1988). That is, equivalence of perceptual structure does not necessarily imply equivalence of performance on tasks that require accurate discrimination and/or recognition of stimuli.

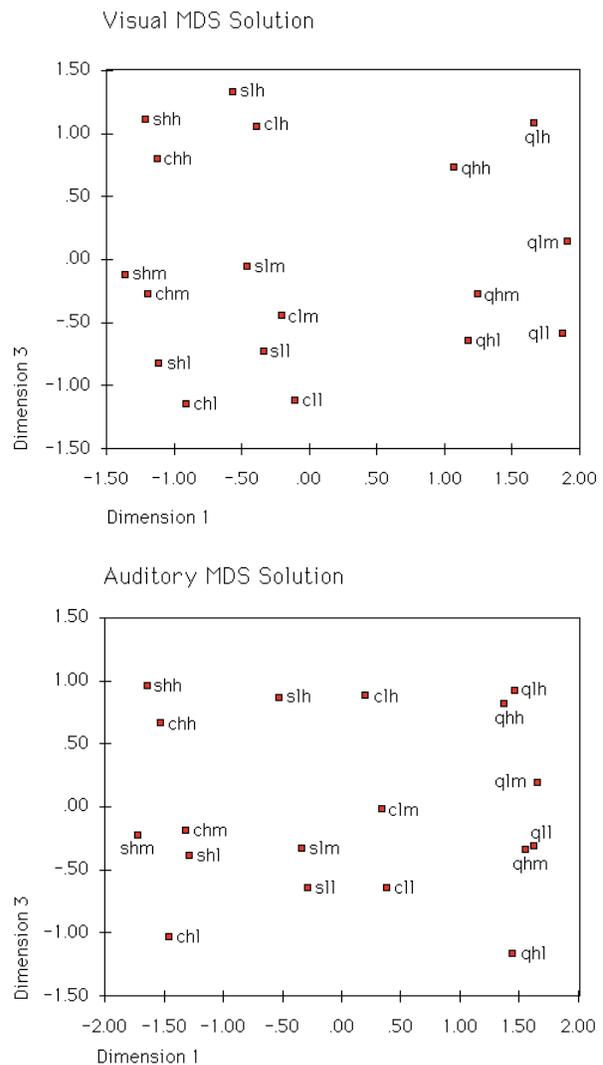


Figure 3. Visual and auditory MDS solutions for Dimensions 1 and 3. Stimuli are denoted according to shape (s, sine; q, square; c, composite), amplitude (h, high; l, low), and frequency (h, high; m, medium; l, low), respectively.

Table 2
Means, Standard Deviations, and Ranges of Percent Correct
on AB-X Discrimination Task for Experiment 2

Dependent Variable	Modality			
	<i>M</i>	<i>SD</i>	Min	Max
Total percent correct				
Vision	95.52	2.88	91.67	100.00
Audition	81.33	9.61	61.11	97.22
A percent correct				
Vision	95.96	4.62	88.24	100.00
Audition	79.10	9.60	58.82	94.12
B percent correct				
Vision	95.05	4.24	88.24	100.00
Audition	83.51	12.26	63.16	100.00

Particularly if auditory displays of graphical stimuli are to be used as a substitute for visual displays, one must consider whether people can discriminate as accurately among similar wave forms when using auditory display as they can when using visual display, regardless of the overall level of perceptual equivalence in representation for the two modalities. When using visual graphs to examine data, people often need to make discriminations between relatively similar graphs. If there is performance equivalence for visual and auditory displays of these wave forms, people should be able to discriminate between similar wave forms equally well with the two types of display. In Experiment 2, we examined the relative performance accuracy of vision and audition on a task involving discriminations between similar wave forms.

Method

Subjects. Thirty-eight undergraduate psychology students who reported normal or corrected-to-normal vision and hearing participated to fulfill a course research requirement. Nineteen subjects (balanced for gender) participated in each modality condition.

Materials. Thirty-six pairs of wave forms selected from same-neighborhood wave forms identified by MDS in Experiment 1 were used as the test stimuli for this experiment. An additional four pairs of wave forms were randomly selected from non-test pairs to serve as stimuli for catch trials.

Procedure. Subjects were assigned to either the visual or the auditory condition to perform a discrimination task. An AB-X matching procedure was employed in which two comparison stimuli were sequentially presented (A and B), followed by a third stimulus (X). The subjects' task was to determine whether X was the same as A, B, or neither. Each subject completed a total of 3 practice and 40 test trials. The subjects were informed prior to the experiment that there would be a very small number of catch trials—that is, trials on which the correct response was “neither.” Four catch trials were included to ensure that subjects attended to both A and B before making their judgments about X, rather than adopting the simpler strategy of ignoring A, attending to B, and making a same-different judgment for the B-X pair (see Garbin, 1988).

In both conditions, the subjects were presented with three practice trials demonstrating the three possible responses in order to familiarize them with the discrimination task. On these practice trials, but not on the test trials, the subjects were provided with feedback for each trial, in the form of the correct answer. Following the practice trials, the subjects completed 40 test trials with a short break at the halfway point.

As in Experiment 1, visual stimuli were presented with an overhead projector. For each trial, each of the three wave forms was

presented only once, for a duration of 4 sec, with a 2-sec interval between wave forms. Auditory stimuli were presented on the computer speaker as in Experiment 1. For each trial, each of the three wave forms was presented only once, for a duration of 6 sec, with a 2-sec interval between wave forms.

Results

Three dependent variables were used to assess subjects' performance on this discrimination task: (1) total percent correct for the 36 test trials (total percent correct), (2) percent correct of the 18 trials on which X was A (A percent correct), and (3) percent correct of the 18 trials on which X was B (B percent correct). On the basis of outlier analysis using these three dependent variables, 2 subjects (one visual and one auditory) were excluded from further analyses.²

Comparisons of A percent correct and B percent correct performance within each modality revealed no significant effect of position of test stimulus on discrimination accuracy ($p > .05$ for all comparisons). These results suggest that the subjects were able to access waves presented first as well as they could access those presented second for the purposes of identifying the X stimulus as either A or B in this AB-X discrimination task. Thus, there appears to have been no effect of memory demand on discrimination performance for either vision or audition.

Analyses of all three dependent variables suggested that visual performance was better than auditory performance [$F(1,34) = 36.03$, $MS_e = 50.36$, $p < .0001$ for total percent correct; $F(1,34) = 45.04$, $MS_e = 56.78$, $p < .0001$ for A percent correct; $F(1,34) = 14.25$, $MS_e = 84.11$, $p < .001$ for B percent correct].³ Table 2 presents means, standard deviations, and ranges of all three variables used for these comparisons.

Although these results suggest that for the purposes of discrimination with these stimuli, audition is inferior to vision, it should be noted that auditory performance was well above chance and that the best performances of auditory subjects were at least as good as the average visual subject performances. Thus, auditory modes of presenting wave-form data should not be completely rejected; the discrimination performance of the best auditory subjects suggests that there may be some manipulation (e.g., practice) that would make auditory performance more equivalent with that of vision.

GENERAL DISCUSSION

The results from Experiment 1 suggest that visual and auditory modes of graphical data presentation result in perceptual representations of those data which are extremely congruent. Although visual and auditory displays resulted in very similar perceptual representations for the two modalities, results from Experiment 2 suggest that these representations were not functionally equivalent; that is, they were not used equally well by the two modalities.

Auditory performance on the discrimination task was well above chance, indicating that this mode of presentation shows promise as an alternative or supplement to

visual graph presentation. However, information represented visually appeared to be superior for the purposes of making detailed, difficult discriminations among these wave forms.

There are at least two possible reasons for the performance differences between the two modalities in Experiment 2. First, the superior performance for visual displays may simply be based on subjects' greater level of experience with visually presented graphs. Particular types of practice in particular amounts may provide subjects with the experience they need in order to perform as well on a discrimination task with auditory as with visual graphs. Second, the auditory displays may not have been scaled for optimal performance. For example, some other scaling factor for mapping amplitude into pitch or for mapping frequency into time may have resulted in improved auditory performance.

Future developments in auditory graphical display should be based on knowledge of how graphs may best be presented and how people can best be prepared to perceive them, for the purposes of making the detailed kinds of discriminations that are currently achieved with more traditional visual displays. Future research, then, should focus on stimulus presentation or subject experience manipulations that could improve the level of auditory discrimination performance to that of vision.

REFERENCES

- 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: Human Perception & Performance*, **14**, 547-553.
- GARBIN, C. P. (1990). Visual-touch perceptual equivalence for shape information in children and adults. *Perception & Psychophysics*, **48**, 271-279.

NOTES

1. Although the R^2 and stress values were acceptable at 2 dimensions for the auditory solution, the 3rd dimension was retained because it was both interpretable and stable. The 3rd dimension was interpreted as a frequency dimension and recovered one of the stimulus attributes "built in" to the stimulus set. To check for stability, the sample of auditory subjects was randomly split into one group of 6 and another of 7 subjects. Composite matrices based on each of the splits were analyzed using MDS, and both revealed R^2 values of .93 and stress values of .09 for a 3-dimensional solution. CC values based on comparisons between the split samples and comparisons of each split sample with the original composite matrix were above .90 for all 3 dimensions.

2. Each of these subjects had extremely low total percent correct and B percent correct scores, as well as a high number of "neither" responses to test trials (where the correct answer was either A or B) compared with the performance of other subjects in their condition.

3. Incorrect responses to test trials could be of two types. Subjects could identify the trial as a test trial but identify X incorrectly (i.e., say "A" when X is actually B), or they could incorrectly identify the trial as a catch trial (i.e., say "neither" when X is actually A or B). When subjects incorrectly identify a test trial as a catch trial, they effectively eliminate any possibility that their answer will be correct. Subjects who correctly identify a test trial at least have the opportunity to make a correct response. Another way of assessing performance for this discrimination task, then, is to correct for trials where subjects incorrectly identify a test trial as a catch trial (thereby eliminating the possibility of a correct response). Analyses were conducted using the same three dependent variables, which were corrected for incorrect identifications of test trials as catch trials. For all three analyses, mean performance levels improved slightly and about equivalently for both modalities relative to performance levels of the uncorrected dependent variables, and the patterns of the effects for modality and target position were identical to those found for the uncorrected dependent variables.

(Manuscript received November 13, 1995;
accepted for publication December 19, 1995.)