Sensory-modality opinion scales for individual subjects

WILLIAM E. DAWSON and MICHAEL A. MIRANDO
University of Notre Dame, Notre Dame, Indiana 46556

In two sessions each, 16 subjects were asked to match apparent handgrip force and apparent duration in a counterbalanced order to the pronounceability of 12 trigrams in order to assess sensory modality opinion scales for individuals. Scales for individuals, like those for group data, were consistent across tasks and showed reliability across sessions. Subjects displayed strong agreement on the relative pronounceabilities of the trigrams according to Kendall's test for concordance. A significant intersession correlation for subjects' multimodality exponents indicated reliable differences in these exponents among individuals. Comparisons with psychophysical scales were made and discussed.

In 1971, Dawson and Brinker showed that direct ratio scales of opinion like those discussed by S. S. Stevens (1966) could also be produced by requesting subjects to match the intensity of their sensory impressions to the intensity of their opinions. Prior to this, ratio scales of opinion were obtained mainly through the use of the method of magnitude estimation which involves the matching of numbers to opinion strength. In contrast, in sensory-modality matching, subjects adjust some physical continuum, e.g., sound pressure, so that its resultant sensory effect, in this case loudness, is in direct proportion to their strength of opinion for each of a series of stimulus items. The procedure is similar to cross-modality matching, but differs in that the items need have no clear physical metric.

Sensory-modality scales have been shown to be consistent with cross-modality matches through an indirect check using multimodality matching. In the latter procedure, subjects match two (or more) sensory modalities to indicate their opinions of the same set of items. Then adjustments of one modality are examined as a function of those of the other for the various items. The function recovered has been found to be similar to that which occurs with cross-modality matching (Dawson & Brinker, 1971).

A sensory-modality scale for the pronounceability of trigrams has been shown to relate to free recall in a memory task at least as well as Underwood and Schulz's (1960) rating scale for this dimension (Dawson & Mirando, 1973). Other attempts to employ sensory-modality matching include the effort of Brennan, Ryan, and Dawson (1975) to scale the perceived accentedness of speech samples. Here, a sensory-modality scale was found to agree with one obtained by magnitude estimation. The relationship

between the two scales was a power function possessing the exponent predicted on the basis of prior psychophysical studies. Scale values, in this preliminary study, related in a clear-cut manner to certain phonological variables.

The foregoing studies indicate that the sensory-modality matching procedure, at least for the data of groups of subjects, results in observer performance that is consistent with performance in psychophysical ratio scaling tasks. They also show that sensory-modality scales can be successfully used in a variety of areas of inquiry. What these early studies have not yet disclosed is whether the performance of individual subjects is like that of the group average as it is for psychophysical scales (e.g., Marks & J. C. Stevens, 1966; J. C. Stevens & Guirao, 1964). Several specific questions need to be answered.

Multimodality matching demonstrated that group data are consistent across two sensory-modality matching tasks: the matches to a set of items for one sensory modality are highly correlated to similar matches for a second sensory modality. Further, the relation between the two sets of matches is describable by a power function having the predicted exponent. But do the data of individual subjects also yield such a cross-modality correlation, and are they, too, fitted by a power function with the appropriate exponent? In the psychophysical domain, individuals appear to conform to the power law although the intersubject variation in exponent is quite high (Marks & J. C. Stevens, 1966; J. C. Stevens & Guirao, 1964).

Another question deals with whether all subjects agree with respect to the various items scaled. The group results purport to scale the items presented along the dimension requested by the experimenter, and presumably all subjects are more or less in agreement about the items. In fact, items are selected with such agreement in mind since scales for stimuli are the goal here rather than scales for individuals (cf. Hays, 1967). However, a significant intermodality correlation plus a power function with the correct

exponent with group data need not indicate intersubject agreement. For example, Subject A may judge an item as very low on the dimension scaled, and he may do so consistently for both sensory-modality matches. Subject B, on the other hand, may judge the same item as high consistently across the two matches. With such variation among the subjects, it is still possible to get the correlation, a power function, and the appropriate exponent. Therefore, some check still needs to be made with regard to subject agreement. Kendall's W is a statistic which can indicate whether there is significant agreement among subjects about the items scaled.

Intersession or repeat reliability also is of concern. Are a subject's adjustments in one sensory-modality matching session closely related to those in a subsequent, second session that involves the same sensory-modality match? Such a relationship occurs for psychophysical scales (e.g., J. C. Stevens & Guirao, 1964). Assuming that nothing occurs between sessions to change subjects' opinions, intersession correlations should be reasonably high, since the same items and the same matching modality are employed. In addition, the intersubject differences in exponents mentioned above also need to be looked for in the exponents recovered with multimodality matching. Thus, if individual differences in exponents are found for a session, do they persist over time as indicated by a significant correlation with a set of such exponents obtained from a later session with the same task?

To answer these questions, subjects were asked to perform two sensory-modality matches to indicate their opinions of the pronounceability of a series of trigrams. Force of handgrip and apparent duration were the continua selected. The two matches were performed in one session and then repeated later in a second session to assess the reliability of the adjustments and of the individual differences.

METHOD

Subjects

Sixteen male undergraduate volunteers who were enrolled in introductory psychology courses served as subjects. All were told that there would be a second session when they volunteered, but none was told what any of the tasks would be. They received extra class credit for participation.

Apparatus

The 12 trigrams, selected to vary in pronounceability, were BAD, CUB, REC, VIZ, POH, YUQ, LTY, RTI, NDR, QZP, RZQ, and YQG. Each was typed in capital letters on a small index card for presentation to the subject. Force of handgrip in kilograms of force was measured on a Lafayette hand dynamometer, which was mounted on a stand that was fastened to the table before which the subject sat. For apparent duration matches, a small white lamp was provided which could be lighted for any length of time desired by the subject merely by pressing down and holding a key. Release of the key turned off the light and stopped a Standard timer which measured the time the lamp had been lighted.

Procedure

Each subject took part in two sessions, separated, on the average, by approximately 2 weeks. Half of the subjects matched force first, and half duration in the first session. During this session, all subjects matched both apparent force and apparent duration to the pronunciation difficulty of each trigram. Each trigram was presented three times during both of these matches. They were presented in an irregular order but with the entire 12 given each time before any further repetition of the series.

In the second session, counterbalancing was again employed. The procedure was exactly like in the first session, except that half of those subjects who had adjusted force first previously did so again, while the remaining half changed to duration. An analogous balance was used with the first-session duration-first subjects.

For each sensory-modality match, the subject was instructed to match the first trigram with any sensory intensity that seemed appropriate. Subsequent trigrams were to be matched in a proportional manner with sensory intensity set in proportion to the subject's opinion of pronunciation difficulty. The first trigrams for each subject were selected from the middle of the range of pronounceability as indicated by earlier studies. For any match, the subject was allowed to use as much time as he wished and to repeat the match until he was pleased with it.

RESULTS

Medians of each subject's matches to each trigram were obtained for his first session, his second session, and the sessions combined. To determine whether adjustments varied across sessions or across trigrams, two analyses of variance were conducted—one for force data and one for duration. In both cases, all medians were first subjected to a logarithmic transformation, since, as with psychophysical data (cf. S. S. Stevens & Guirao, 1962), responses were log normal in distribution. Log medians for the two sessions were taken together, and within-subjects, two-factor analyses were run. With handgrip force, the trigrams yielded significantly different squeezes, $F(11,165) = 62.86$, $p < .001$, but the difference between sessions, $F(1,15) = 2.64$, and the Trigrams by Sessions interaction, $F(11,165) = 0.69$, were not significant. Similar results were found with duration: trigrams made a difference, $F(11,165) = 93.81$, $p < .001$, while the sessions effect, $F(1,15) = 1.46$, and the interaction, $F(11,165) = 0.92$, were nonsignificant.

Next, the group data were examined to see if the usual relationships were obtained. The multimodality matching results are presented in Figure 1 for both sessions separately and for their pooled combination. For each of the trigrams, median force (F) is plotted as a function of median time (T) in log-log coordinates. The best fitting straight lines for the three sets of data were found using a method described by Kenney (1939). His procedure locates the best line in a case like the present where both variables are subject to error. Best fitting equations are $\log F = 0.70 \log T - 0.13$ for Session 1, $\log F = 0.73 \log T - 0.22$ for Session 2, and $\log F = 0.71 \log T - 0.17$ for the combined data.

The linearity with log-log axes indicates that a power function describes the relation between the two sensory modality matches in all three cases. The fitted slopes correspond to power function exponents which are close in magnitude to the value expected on the basis of the exponents previously reported for handgrip force, 1.7, and apparent duration, 1.1 (Stevens, 1961). The predicted slope is 1.1/1.7, or 0.65, a value only slightly lower than the slope obtained.

Cross-modality correlations for group data, as would be expected from the above findings, were also quite large. These correlations were calculated between median force and median duration over the 12 trigrams. Pearson rs were 0.98, 0.99, and 0.99 for Session 1, Session 2, and sessions combined, respectively. Clearly, the subjects as a group were able to indicate reliably their opinions using two different sensory modalities.

Table 1 summarizes the analyses of the data for individual subjects. It includes the recovered multimodality matching exponents, the cross-modality correlations for the three types of session analyses, and, in addition, the intersession or repeat reliabilities for both of the sensory-modality matching tasks. The latter two columns of correlations show that subjects are consistent in their matches from one session to the next. All but one of these correlations, namely 0.41, are significant also (p < .01), showing that individuals, as well as the group, can reliably assert their views utilizing different sensory modalities.

The exponents for individuals are highly variable—most subjects' exponents depart consider-
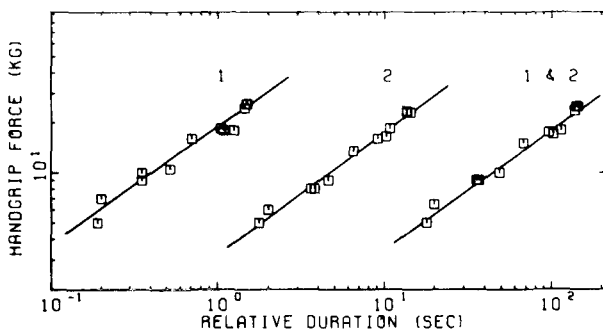

Figure 1. Results of the multimodality match for the first and second sessions and for the data combined. The latter two curves have been displaced one and two log units, respectively, to the right.

ably from the predicted 0.65, even though the exponent for the group data is near this value. Log-log multimodality matching plots for individuals fitted by the Kenney procedure show that the data are consistent with a power function but that the variation about the lines of best fit is considerable. Figure 2 displays eight such plots for eight subjects selected at random. They are based on subjects' data combined across sessions. The differences in slopes and the variation about the fitted lines are apparent. These differences in subjects' exponents are reliable over time, as indicated by a significant intersession correlation, r = 0.64, p < .01 (cf. Rule, 1966).

Although there are subject differences in the matches made to the trigrams, the subjects are in good agreement concerning the relative pronounceability of the various trigrams. Agreement was assessed by arranging the subjects' median matches in Subjects by Trigrams matrices and computing

Table 1
Multimodality Exponents, Cross-Modality Correlations, and Intersession Reliabilities for Individual Subjects

| Subject | Multimodality Exponent | | | Cross-Modality Correlation | | | Intersession Reliability | |
	Session 1	Session 2	Combined	Session 1	Session 2	Combined	Force Adjusted	Duration Adjusted
T.A.R.	.92	.57	.85	.93	.90	.94	.96	.92
P.S.	.33	.46	.36	.92	.90	.94	.90	.83
B.M.	.57	.81	.65	.94	.93	.99	.92	.95
D.W.	.90	.92	.87	.89	.87	.92	.81	.97
B.P.	1.06	1.04	.99	.84	.92	.98	.69	.86
J.G.	1.55	1.47	1.56	.86	.85	.90	.95	.99
T.R.	1.55	1.73	1.31	.77	.92	.89	.41	.95
P.M.	.48	.55	.54	.84	.95	.93	.83	.96
B.N.	.28	.35	.31	.87	.83	.83	.94	.85
J.C.U.	.58	1.26	.94	.89	.97	.97	.88	.89
M.W.	1.00	.65	.95	.86	.83	.94	.84	.80
G.S.	1.06	.55	.83	.96	.84	.96	.91	.93
R.S.	1.02	1.76	1.47	.95	.98	.98	.96	.93
J.C.L.	1.81	1.06	1.37	.92	.96	.98	.94	.95
F.A.	.44	.52	.50	.93	.90	.97	.90	.96
P.F.	.57	.70	.64	.96	.95	.99	.95	.96
Mean	.883	.900	.884	.896	.906	.944	.862	.919
SD	.459	.450	.385	.053	.050	.044	.140	.056
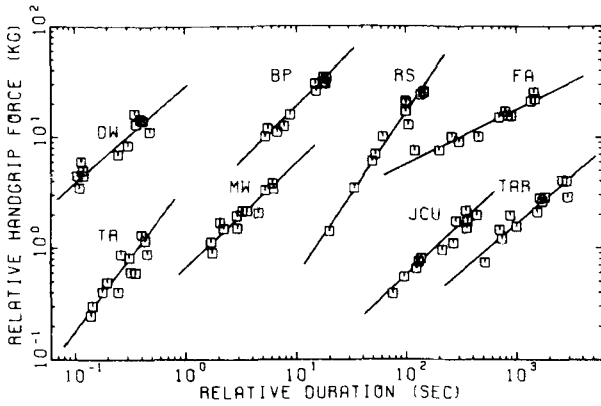
Figure 2. Multimodality matches for eight individuals. Coordinates are correct for Subject D.W.; data for others in the top row are displaced successively one log unit to the right. Data in the bottom row are similar but were displaced downward a log unit also.

Kendall's coefficient of concordance (W). As Table 2 shows, Ws were significant for force and for duration adjustments in both sessions and for the pooled-sessions data as well.

DISCUSSION

Data for the group agree with the reported findings of Dawson and Brinker (1971) in recovering a power function from the multimodality matches and in finding a high cross-modality correlation. And, as they found, the size of the recovered exponent is of approximately the size predicted on the basis of earlier psychophysical research.

The results for individual subjects show that the group function is not merely an artifact of averaging (cf. Jones & Marcus, 1961); the data of individuals are also describable by power functions. Graphical plots of the data of individuals resemble those found for the psychophysical scales of loudness (J. C. Stevens & Guirao, 1964) and brightness (Marks & J. C. Stevens, 1966). The variation in subjects' exponents for opinions, however, may be considerably greater. For example, the coefficients of variation of the individual exponents reported by J. C. Stevens and Guirao (1964) are 0.29 for their first session and 0.22 for their second, as compared to 0.52 and 0.50 for the present data. There are many different factors which could account for this difference, but note that some of the

Table 2
Kendall's W for the Subjects by Trigrams Matrices

Adjusted Continuum	Session 1	Session 2	Combined
Force	.84	.86	.88
Time	.88	.88	.90

*p < .001

additional variation is probably due to the fact that the opinion scale exponents are based on *two* sensory-modality matches, and hence some increase in variability is expected.

Although subjects differ much one from another, they show little change across the two sessions and are quite reliable in their judgments. Further, though they differ in how they express their opinions, individuals are consistent in their reports even with a change in sensory modalities. In addition, all of the subjects are in agreement in their assessments of the relative pronounceabilities of the trigrams—a necessity, if we wish to claim that the scale values represent the opinions of individuals and are not just an artifactual result of the averaging process.

The present study joins many others (e.g., Rule, 1966; Rule & Markley, 1971; Wanschura & Dawson, 1974) in finding that the individual differences in exponents persist over time. Such studies report significant intersession correlations for individual exponents for the psychophysical case. Opinion scales obtained by direct ratio methods appear to be no different in this regard, even though we are looking at exponents recovered by multimodality matching. The source of such differences is still in dispute (cf. Engeland & Dawson, 1974), but the present results do appear to deny S. S. Stevens' (1971) suggestion that they may be largely due to the "regression effect."

The multimodality matching procedure presumably results in an exponent that is free of regression influences. As S. S. Stevens (1971) has stated, when two sensory continua, A and B, are matched to a common sensory continuum, C, "the two matching functions provide exponents whose ratio determines the exponent of the power function relating A to B," and this may "cancel a major portion of the regression bias." If the two sensory-modality matches are seen as analogous to A and B, and if C is replaced by the continuum of pronounceability, then the multimodality matching case is seen as similar to the one Stevens describes. But, still, strong individual differences were found in the present study. Comparable postregression-balance results occurred with cross-modality matching tasks in an experiment by Wanschura and Dawson (1974). These two studies, one with opinion scales and the other with psychophysical scales, demonstrate that regression is not the main or sole source of idiosyncratic exponents.

In summary, the results for individual opinion scales obtained by sensory-modality matching provide further evidence of a parallel between such scales and the psychophysical scales achieved by analogous methods. Magnitude estimation, cross-modality matching, and sensory-modality matching scales share much in common—probably even the same problems. For example, magnitude estimation scales of opinion are nonlinearly related to category scales for the same items (S. S. Stevens, 1966).

REFERENCES

BRENNAN, E. M., RYAN, E. B., & DAWSON, W. E. Scaling of apparent accentedness by magnitude estimation and sensory modality matching. *Journal of Psycholinguistic Research*, 1975, **4**, 27-36.

DAWSON, W. E., & BRINKER, R. P. Validation of ratio scales of opinion by multimodality matching. *Perception & Psychophysics*, 1971, **9**, 413-417.

DAWSON, W. E., & MIRANDO, M. A. Sensory-modality scale for pronounceability of trigrams and its relation to free-recall learning. *Perception and Motor Skills*, 1973, **36**, 1219-1224.

ENGELAND, W., & DAWSON, W. E. Individual differences in power functions for a one-week intersession interval. *Perception & Psychophysics*, 1974, **15**, 349-352.

HAYS, W. L. *Quantification in psychology*. Belmont, Calif: Brooks/Cole, 1967.

JONES, F. N., & MARCUS, M. J. The subject effect in judgments of subjective magnitude. *Journal of Experimental Psychology*, 1961, **61**, 40-44.

KENNEY, J. F. *Mathematics of statistics*. Part 1. New York: Van Nostrand, 1939.

MARKS, L., & STEVENS, J. C. Individual brightness functions. *Perception & Psychophysics*, 1966, **1**, 17-24.

RULE, S. Subject differences in exponent of psychophysical power functions. *Perceptual and Motor Skills*, 1966, **23**, 1125-1126.

RULE, S., & MARKLEY, R. P. Subject differences in cross-modality matching. *Perception & Psychophysics*, 1971, **9**, 115-117.

STEVENS, J. C., & GUIRAO, M. Individual loudness functions. *Journal of the Acoustical Society of America*, 1964, **36**, 210-213.

STEVENS, S. S. The psychophysics of sensory function. In W. A. Rosenblith (Ed.), *Sensory communication*. New York: Wiley, 1961.

STEVENS, S. S. A metric for the social consensus. *Science*, 1966, **151**, 530-541.

STEVENS, S. S. Issues in psychophysical measurement. *Psychological Review*, 1971, **78**, 426-450.

STEVENS, S. S., & GUIRAO, M. Loudness, reciprocality, and partition scales. *Journal of the Acoustical Society of America*, 1962, **34**, 1466-1471.

UNDERWOOD, B. J., & SCHULZ, R. W. *Meaningfulness and verbal learning*. New York: Lippincott, 1960.

WANSCHURA, R. G., & DAWSON, W. E. Regression effect and individual power functions over sessions. *Journal of Experimental Psychology*, 1974, **102**, 806-812.