Validation of ratio scales of opinion by multimodality matching

WILLIAM E. DAWSON and RICHARD P. BRINKER
University of Notre Dame, Notre Dame, Indiana 46556

Ss adjusted loudness and force of handgrip (and in one experiment apparent duration) to match the strength of their opinions regarding racism, occupational desirability, and pronunciability of trigrams. For each stimulus item, median settings of one continuum were plotted as a function of median settings of the other. The resulting power-function exponents agree with previous cross-modality matches and with the results of a cross-modality match of loudness and duration produced by Ss in the present experiments. The results provide validation of ratio scales of opinion and further establish the direct scaling methods used to obtain them.

S. S. Stevens (1966a, 1968) has summarized experiments from several laboratories that suggest that procedures developed to scale sensory attributes may be used to measure opinions and attitudes. Direct scaling methods, e.g., magnitude estimation and ratio estimation, have been used to scale such variables as utility, occupational preference, esthetic value, and seriousness of offenses. As Stevens points out, scales obtained by these methods often stand in invariant relations to scales for the same variables obtained by other methods. Pair comparison and category confusion scales are logarithmic functions of the corresponding ratio scale. A simple category scale plotted as a function of the corresponding ratio scale yields a curve that is concave downward. The occurrence of these invariant relations with attitudinal continua parallels many of the sensory continua for which the same invariant relations have been found (e.g., Galanter & Messick, 1961; Stevens & Galanter, 1957). The recurrence of the same interscale relations suggests that direct scaling methods can be used to produce ratio scales of opinions and attitudes.

Ratio scales for sensory continua have been validated by direct cross-modality matching. In that procedure Ss adjust the sensory intensity of one modality to match in apparent intensity various sensory intensities of another modality. Since ratio-scaling methods find that sensory intensity is related to physical intensity by a power function (e.g., see Stevens, 1961), the result of a cross-modality match can be predicted. For the two sense modalities, sensory intensity is given by the equations

$$\psi_1 = \phi_1{}^a \qquad (1)$$

and

$$\psi_2 = \phi_2{}^\beta \qquad (2)$$

where ψ_1 and ψ_2 are the sensory intensities equated at various levels, ϕ_1 and ϕ_2 are the associated physical intensities, and a and β are the exponents characteristic of the two sensory modalities. In cross-modality matches ψ_1 is set equal to ψ_2; hence,

$$\phi_1{}^a = \phi_2{}^\beta \qquad (3)$$

or in logarithms

$$\log \phi_1 = (\beta/a)(\log \phi_2). \qquad (4)$$

Thus in log-log coordinates the equal-sensation function is a straight line having a slope β/a. A variety of such predictions of linearity with specified slope have been confirmed by experiment (Stevens, 1961).

Unlike psychophysical continua, there are no power functions that describe the relation between intensity of opinion and some physical dimension since no obvious physical continuum is involved. Nevertheless, a similar kind of validation may be attempted. Ss can be asked to indicate the strength of their opinion about some variable by adjusting the intensity of a physical stimulus so that the resultant sensory intensity matches the strength of their opinion. For validation the S must match two (or more) such sensory magnitudes to his intensity of opinion. Each of the two matches would presumably result in a ratio scale of opinion for the given set of stimuli (see Stevens, 1966b). That is, to any stimulus θ_i of a set of n stimuli there would correspond two sensory intensities, ψ_{1i} and ψ_{2i}. If stability of opinion is assumed, then the opinion of θ_i will remain the same when matched to both sensory continua. In fact, ψ_{1i} will equal ψ_{2i} since things matched as equal to the same thing should be equal. Thus, for each of the n stimuli there will be a pair of equal sensory intensities. But ψ_{1i} and ψ_{2i} are in turn related to their associated physical continua by the power functions given in

Eqs. 1 and 2. And since ψ_{1i} and ψ_{2i} are presumed equal, Eqs. 3 and 4 hold once again.

What the foregoing means is that if the n pairs of θ_{1i} and θ_{2i} are plotted in log-log coordinates, the result should be a straight line with a slope of β/a. That is to say, the same result is obtained here as is obtained in direct cross-modality matches between the two sensory continua.

The experiments to be reported address the questions: (1) Can Ss indicate the intensity of their opinions by matching them to sensory intensities produced by varying the level of a physical stimulus? (2) When Ss attempt to make such matches using two different types of physical stimuli, are their matches in accord with the experimental results of a cross-modality match that involves these same physical stimuli and their corresponding sensations?

Table 1
Behaviors Scaled for Racism

| Code Designation	Behavior
1	A group of white men hanged and mutilated a Negro man who was observed talking to a young white girl on the street.
2	A white man joined several of his neighbors in setting fire to a home owned by a Negro family, the first such family to move into the man's neighborhood.
3	A white man joined the Klu Klux Klan.
4	The white owner of a small business company who had hired the man doing the best on a competitive examination for a job fired him when he saw that he was a Negro.
5	A white administrator of a southern university refused to admit Negro students.
6	The white owner of a motel refused to rent a room to a Negro family that was traveling.
7	A white guest at a small party loudly criticized the hostess for inviting a Negro couple and immediately left the party.
8	A white man gave a speech urging that segregation be enforced in order that neither of the two races would contaminate one another.
9	A white woman refused to send her son and daughter to a school when some Negro children were permitted to enroll.
10	A white judge upheld a state law that prevented a Negro man from marrying a white girl.
11	A white man said that he couldn't understand why Negroes complained—they had as much of an opportunity as anyone else to get ahead in life.
12	A white girl avoided social gatherings at which the majority present were Negroes.
13	A white student stated that he felt uneasy around black militants.
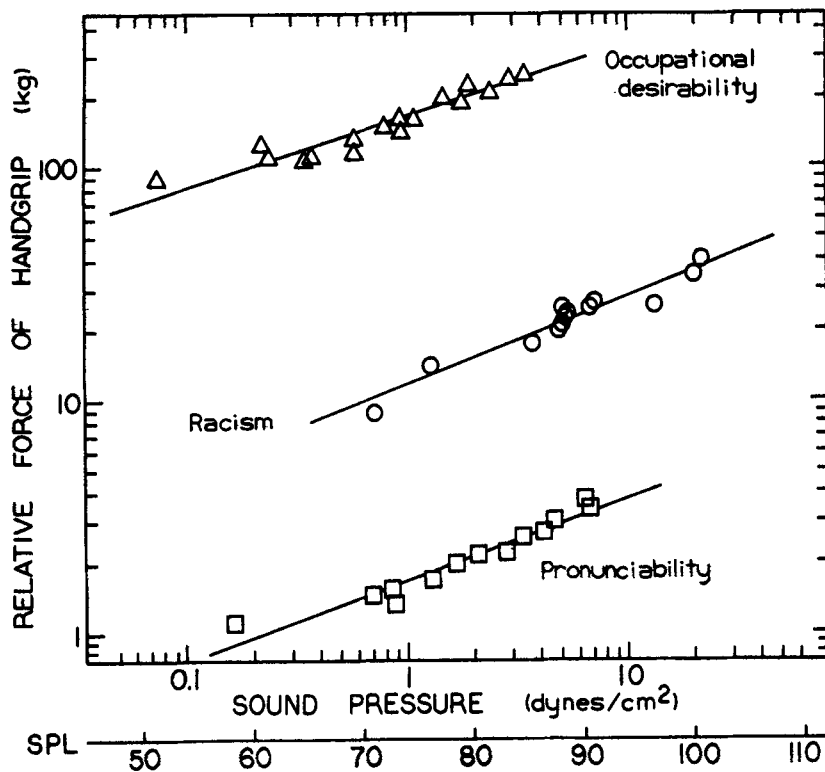
Fig. 1. Results of the multimodality matches in Experiments 1 and 2. The top curve has been shifted upward one log unit and the bottom curve, downward one log unit.

EXPERIMENT 1:
RACISM

Ss read a series of 13 sentences, each of which described some behavior on the part of a white person. An S's task was to indicate his opinion of how racist was the behavior mentioned by adjusting the loudness of a tone and by squeezing a hand dynamometer.

Subjects

The Ss were 24 male undergraduates from the University of Notre Dame. Most of them were enrolled in introductory psychology courses and received extra credit for participation.

Apparatus

The 13 sentences, which were each typed on a single index card, are listed in Table 1. A Lafayette hand dynamometer was used to record force of handgrip in kilograms. An audio oscillator provided a 1,000-Hz tone whose output level was controlled by a variable attenuator and a sone potentiometer placed in series with it. The S listened to the tone with a pair of earphones so calibrated that sound pressure readings in dynes/square centimeter could be calculated from the voltage measured across them.

Procedure

The sentences were presented one at a time to S in an irregular order that varied from S to S. Each statement was presented twice to each S. Half of the Ss first matched apparent loudness to the degree of racism exemplified by each of the behaviors. The S made this match by adjusting the sone potentiometer to give the appropriate loudness. The E adjusted the variable attenuator to ensure that the full range of desired loudnesses was available to S. Next this group of Ss matched the apparent force of handgrip to their subjective impression of the amount of racism. The remaining half of the Ss first matched apparent force and then apparent loudness to the perceived intensity of racism. At the outset of each matching task, S read a set of instructions that asked him to make matches that were in proportion to the intensity of his opinion. For the first sentence, S was to make whatever match seemed to him appropriate. Then if a subsequent behavior appeared twice as racist, S was to set twice the loudness or apparent force, and so forth.

Results

The median handgrip force (kilograms) and the median sound pressure (dynes/square centimeter) were calculated for each of the 13 sentences. In Fig. 1, median force is shown in a log-log plot as a function of the corresponding median sound pressure for each of the statements. Equivalent sound-pressure levels (SPL re 0.0002 dyne/cm^2) are also given along the abscissa. The method of least squares was used to obtain the following best-fitting straight line for these data: $\log F = 0.39 \log P + \log 11.5$. In this equation F refers to force and P to sound pressure. Taking the antilogarithm of both sides gives $F = 11.5\, P^{0.39}$, which suggests that force of handgrip is related to sound pressure by a power function, with 0.39 as the exponent.

The reported exponent for force of handgrip is 1.7 (Stevens & Mack, 1959); however, an exponent of 2.0 is reported for some methods, such as magnitude production. The loudness exponent is reported as 0.60 (Stevens, 1955), as 0.64 (Stevens, 1966c), and may be as high as 0.68, according to data reported by Stevens (1969). For purposes of comparison and discussion, values of 1.7 for force of handgrip and 0.64 for loudness will be assumed. Thus the predicted slope for Fig. 1 is 0.64/1.7, or 0.38, a value quite close to the 0.39 obtained experimentally.

The experiment suggests, therefore, that persons can indicate the strengths of their opinions by matching sensory intensities to them. Furthermore, Ss seem able to match sensory intensities from two different sensory continua in a manner that is consistent with the previously reported results of a cross-modality match of these same two sensory continua.

A check was also made on the degree of agreement between the levels of sound pressure and handgrip force produced for the statements. A rank-order correlation of 0.97 was found between median sound pressure and median handgrip force. Such a high "cross-modality reliability" indicates that Ss are consistent in their settings for the two physical continua.

EXPERIMENT 2:
OCCUPATIONAL DESIRABILITY
AND PRONUNCIABILITY

Opinion scaling was now carried out for two new attitudinal continua. Ss expressed their opinions on the desirability of certain occupations and also on the *ease* of pronunciability of selected trigrams. The S's task was again to indicate the strength of his opinion by adjusting the loudness of a tone and by squeezing a hand dynamometer.

Subjects

The Ss were 24 male undergraduates from the University of Notre Dame. Again, most were enrolled in a course in introductory psychology, though none had taken part in Experiment 1.

414

LOUDNESS   BRIGHTNESS   RACISM   OCCUPATIONAL DESIRABILITY   PRONUNCIABILITY
Exp. 2   Exp. 3

$\dfrac{20.0 \ \text{dynes}}{\text{cm}^2}$ — 64 sones 2.97×10^{-1} lambert — 64 brils   1 — 2 ; 3

6.32 — 32 3.64×10^{-2} — 32   5,7,9,11 — 4,6,8,10   CUB — BAD YQG ; REC ; JUX

Physician — College professor

2.00 — 16 4.45×10^{-3} — 16   High school teacher — Business executive   IFO — POH TLY ; Architect — Engineer   TLY — YUQ   RTI

12   Photographer — Salesman ; Artist — Forestry officer   LTY — GVS ; RTI

0.632 — 8 5.45×10^{-4} — 8   13   Farmer — Cashier   VIZ — IFO

Sailor — Farm hand

0.200 — 4 6.68×10^{-5} — 4   Bus driver — Factory worker   YQG — BAD

0.0632 — 2 8.18×10^{-6} — 2   Hairdresser

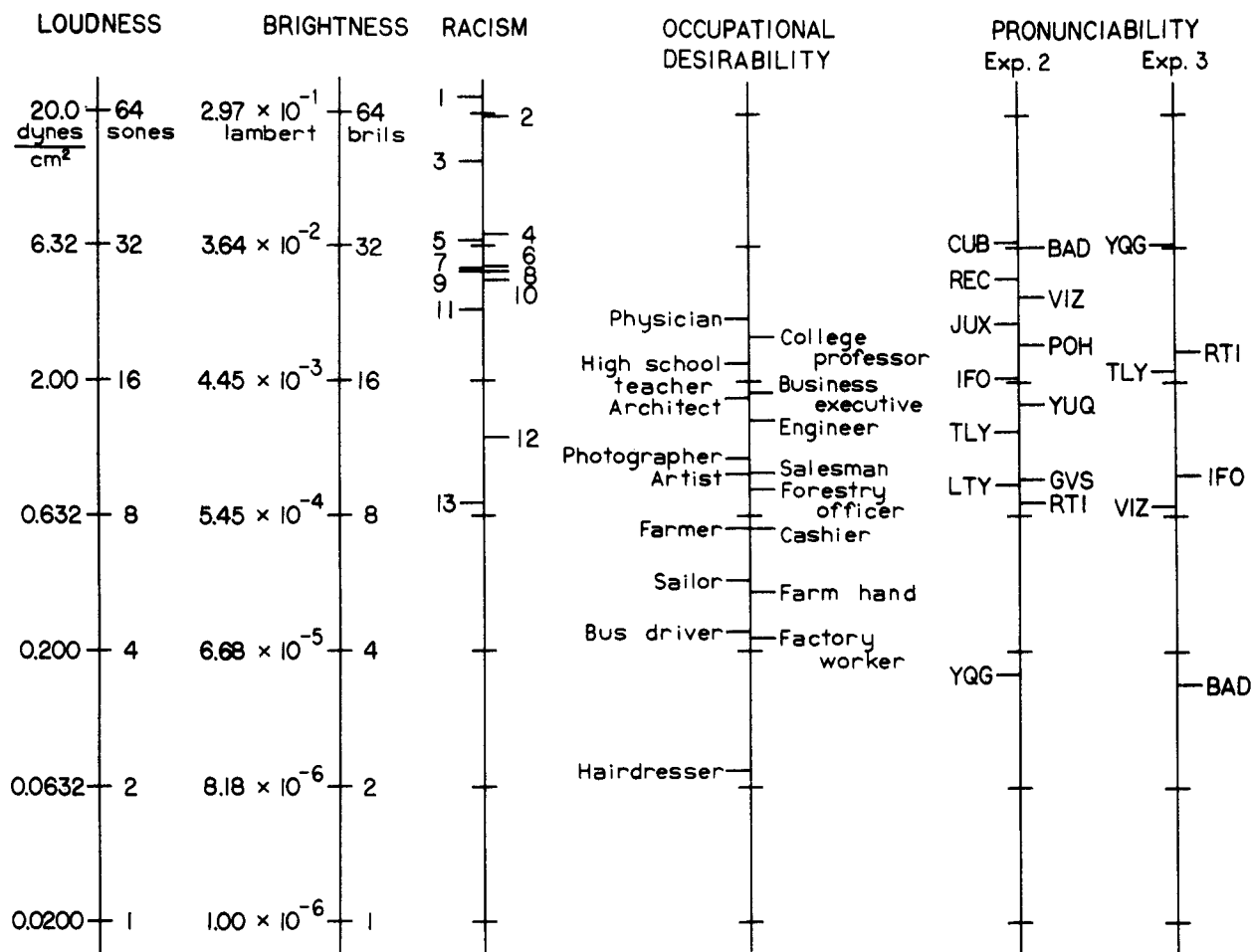0.0200 — 1 1.00×10^{-6} — 1

Fig. 2. A comparison of ratio scales for two sensory continua and three attitudinal continua. The reference scale is the sone scale of loudness. The bril scale has been aligned with it such that one bril equals one sone. Scale position on the opinion scales was determined by the sound pressure matched to the stimulus as shown along the loudness scale. Numbers on the racism scale refer to the behaviors listed in Table 1.

Apparatus

The occupations selected for scaling were taken from Künnapas and Wikström (1963). However, 3 of their 17 occupations were changed. "Bus conductor" was changed to "bus driver" since Americans seldom encounter a bus conductor. "Painter" was felt to be ambiguous; for example, did it refer to a house painter or an artist? "Manufacturer" was changed for similar reasons. In place of these two occupations, "business executive" and "college professor" were added in an attempt to provide more occupations at the high end of the scale. Each of the 17 occupations was typed on an index card. The occupations used are presented in Fig. 2.

The trigrams to be scaled for pronunciability were taken from a list by Underwood and Schulz (1960, pp. 370-372). They constructed their list by having Ss rate, on a 9-point scale, the pronunciability of many three-letter sequences. The 13 trigrams selected from the list were more or less evenly distributed between ratings of 1 ("easy") and 9 ("hard"). Each of the trigrams was typed in capital letters on a single index card. They are also listed in Fig. 2.

The apparatus described in Experiment 1 was used to obtain subjective impressions of occupational desirability and trigram pronunciability. Each S again varied both loudness of tone and apparent force of handgrip to indicate his opinion.

Procedure

Half the Ss were given the occupational stimuli first and the trigrams second. The other half received the trigrams first. Each half was further divided in half: one group first matched apparent loudness in proportion to strength of opinion and then apparent force of handgrip; the other group first matched apparent force. Each S was required to produce two matches to each stimulus card. The complete set of cards for occupations (or trigrams) was given to S one at a time in an irregular order and then repeated, but in a different irregular order.

For occupational desirability the instructions read, "Your task is to indicate how much you would like the occupation to be your permanent occupation." For the trigrams the instructions read, "Your task is to indicate how easy the sequence is to pronounce."

Results

The data for the two tasks were treated in the same way as the data in Experiment 1. The results of the analyses are presented in Fig. 1, along with the straight lines fitted to the data by the method of least squares. In both cases the datum points tend to fall along a straight line in the log-log plot. The slope of the line for the data that involved occupations is 0.29; that for the trigrams is 0.34. The value for the trigrams is fairly close to the predicted slope of 0.38, but 0.29 departs somewhat more from this value.
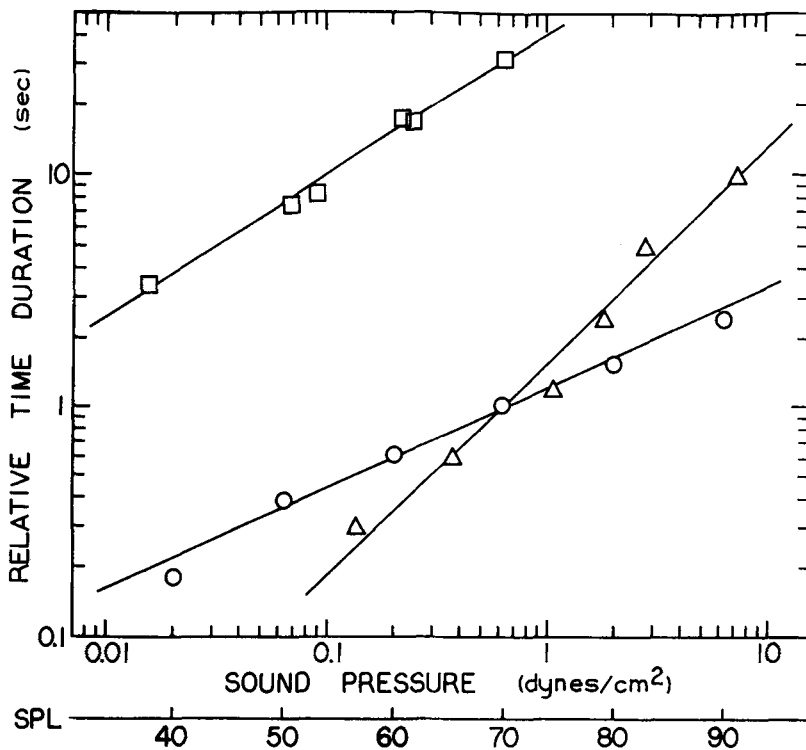
Cross-modality reliabilities were again

Fig. 3. Results of the multimodality match (top curve) and the cross-modality matches (bottom curves) in Experiment 3. Circles indicate that Ss adjusted time duration; triangles indicate that sound pressure was adjusted. The top curve has been shifted both upward and to the left by one log unit.

calculated. A rank-order correlation of 0.94 was obtained for occupations and 0.98 for trigrams.

EXPERIMENT 3: COMPARISON WITH CROSS-MODALITY MATCHING

The next step was to see how the multimodality matching procedure used in the first two experiments would compare with *direct* cross-modality matches made by the same Ss. The Ss again matched the levels of two sensory continua to the intensity of their opinions. In addition, all Ss performed two cross-modality matches: First, levels of one sensory continuum were matched to levels of the other and then vice versa. The equation relating the two continua was then compared with the equation based on the multimodality matches to opinion.

Subjects

The Ss were 24 undergraduates (one female) from introductory classes in psychology. Participation was optional, but class credit was given to those who took part. Three of the Ss had also participated in Experiment 1.

Apparatus

Six of the trigrams used in Experiment 2

were selected to be scaled for pronunciability (see Fig. 2 for those chosen). The apparatus described in Experiment 1 was used to produce a 1,000-Hz tone whose sound-pressure level could be varied as required. Values of 40, 50, 60, 70, 80, and 90 dB were presented to S in the cross-modality matching task.

Apparent time duration was selected as the second sensory continuum because fixed time intervals could easily be presented and repeated. Various durations were produced by the lighting of a small lamp for specific periods of time. For the multimodality match and one cross-modality match, S caused the lamp to light by pressing a key. Release of the key turned off the lamp and a Standard timer recorded the length of time the lamp was lighted. During the other cross-modality match, a Hunter timer was used to present time intervals of 0.3, 0.6, 1.2, 2.4, 4.8, and 9.6 sec.

Procedure

Half the Ss performed the multimodality match to the trigrams first. Of those Ss one-half first matched apparent loudness to pronunciability and then apparent duration; the others made matches in the reverse order.

The other half of all the Ss did the

cross-modality match first. They were divided into halves also: one group first matched loudness to duration and then duration to loudness; the other group matched in the reverse order.

After completion of the multimodality matches, a S commenced the cross-modality matches, or vice versa. All Ss made all matches, and all matches were performed during a single session. For each of the tasks, the six stimuli (trigrams, tones, or time intervals) were presented in an irregular order. Each S made each type of match to each stimulus twice before he began a new task.

In the experiment with the trigrams, Ss were now instructed to indicate how *difficult* they were to pronounce. The S either adjusted the sone potentiometer to give the matching loudness or operated the key to light the lamp for the appropriate duration.

For the first stimulus presented in the cross-modality matches, S was told to make whatever match seemed to him appropriate. Then if a subsequent stimulus produced a sensory impression twice as intense, S was to set twice the loudness (or apparent duration), and so forth. His matches were to be in proportion to his sensory impressions.

Results

The median sound pressure and median time duration matched to each of the trigrams were determined. In the top curve of Fig. 3, median time duration is plotted as a function of the corresponding median sound pressure for each of the trigrams. A straight line with a slope of 0.60 has been fitted to the data by the method of least squares. Since the reported exponent for apparent duration is 1.1 (Stevens, 1961), the predicted slope for this line is 0.64/1.1, or 0.58, a value only slightly lower than that obtained.

The results of the two cross-modality matches were also analyzed; median sound pressures and median time durations produced by the Ss were determined for each of the stimuli presented for matching. The bottom two curves in Fig. 3 present these medians as a function of the presented stimulus (sound pressure or time interval). The circles indicate the matches of duration made to the various tones; the triangles indicate the median sound pressures set to the given time intervals. Straight lines in the log-log plot provide a good fit to these equal-sensation functions. The slopes of the best-fitting lines, as determined by the method of least squares, are 0.44 and 0.91.

It has been noted by other Es that two different slopes are obtained when two sensory continua are matched in this

manner. The O tends to restrict the range of his adjustments on whichever variable he is asked to control. Stevens and Greenbaum (1966) cite a variety of experiments in which this "regression effect" has been found. Indow and Stevens (1966) have suggested that taking the geometric mean of the two slopes might be an appropriate way of averaging out the regression effect. In the present case, the geometric mean of 0.44 and 0.91 is 0.63. Note that one of the slopes is flatter and the other steeper than the slope of 0.60 found with the opinion judgments, but that the geometric mean of the slopes, 0.63, is almost the same as that value.

The cross-modality reliability was again quite high. A rank-order correlation of 0.94 was obtained between the sound pressures and time intervals produced for like trigrams.

DISCUSSION

The experiments provide further validation of ratio scales of opinion and the direct scaling procedures on which they are based. Previous ratio scaling of opinions has generally involved the matching of numbers to intensities of opinion. The present results show that Ss can also adjust sensory intensities to indicate the strength of their opinions.

When such matches are made for two sensory continua, they are done in a manner consistent with the results of a direct cross-modality match of the two sensory continua. That is to say, the power-function exponent obtained from the cross-modality match agrees with the exponent found when the matches for the one sensory continuum are plotted against the corresponding matches for the other sensory continuum. In the past, cross-modality matches have been used to validate the matching of numbers to sensory experience; multimodality matches provide a way to validate the matching of

numbers (or any other continuum) to opinions.

The direct scaling methods, it appears, can be used to scale a wide variety of variables: sensory impressions, opinions, and probably other sorts of judgments. Sensory psychophysics and opinion scaling are thereby brought together within a framework that may provide a basis for the quantification of the full gamut of sensory experience, opinion, and judgment. Figure 2, a small step in this direction, presents the scales obtained in the present experiments, alongside the loudness (sone) and brightness (bril) scales found by Stevens. The primary scale in relation to which the others have been drawn is the sone scale (with exponent equal to 0.60).

Although the experimentally obtained exponents are not identical to the predicted ones, it is likely that the agreement will be adequate when the variability of such exponents is taken into account. More information on exponent variability than is currently available is needed to assess the degree of agreement. Stevens (1969) discusses the 68 experiments of Moskowitz that involved matches of numbers and of loudnesses to various levels of taste intensity. When magnitude estimations were plotted as a function of sound pressure for their common respective levels of taste intensity, 68 estimates of the loudness exponent were obtained. The distribution of exponents had a geometric mean of 0.68 and a standard deviation of 0.51 decilog unit. The multimodality matching of loudness and apparent force of handgrip would yield its own distribution of exponents, as would that for loudness and apparent duration. The Moskowitz data are cited only to give some idea of what sort of distributions might be obtained. More precise measures of exponents and their variabilities will become available as more and more data are accumulated using the

multimodality method to obtain "regression-free" estimates of exponents. Comparisons between predicted and obtained exponents will likewise become more exact.

REFERENCES

GALANTER, E., & MESSICK, S. The relation between category and magnitude scales of loudness. Psychological Review, 1961, 38, 363-372.

INDOW, T., & STEVENS, S. S. Scaling of saturation and hue. Perception & Psychophysics, 1966, 1, 253-271.

KÜNNAPAS, T., & WIKSTRÖM, I. Measurement of occupational preferences: A comparison of scaling methods. Perceptual & Motor Skills, 1963, 17, 611-624.

STEVENS, J. C., & MACK, J. D. Scales of apparent force. Journal of Experimental Psychology, 1959, 58, 405-413.

STEVENS, S. S. The measurement of loudness. Journal of the Acoustical Society of America, 1955, 27, 815-829.

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, 1966a, 151, 530-541.

STEVENS, S. S. On the operation known as judgment. American Scientist, 1966b, 54, 385-401.

STEVENS, S. S. Matching functions between loudness and ten other continua. Perception & Psychophysics, 1966c, 1, 5-8.

STEVENS, S. S. Ratio scales of opinion. In D. K. Whitla (Ed.), Handbook of measurement and assessment in behavioral sciences. Reading, Mass: Addison-Wesley, 1968.

STEVENS, S. S. On predicting exponents for cross-modality matches. Perception & Psychophysics, 1969, 6, 251-256.

STEVENS, S. S., & GALANTER, E. Ratio scales and category scales for a dozen perceptual continua. Journal of Experimental Psychology, 1957, 54, 377-411.

STEVENS, S. S., & GREENBAUM, H. B. Regression effect in psychophysical judgment. Perception & Psychophysics, 1966, 1, 439-446.

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