

Judgments of differences and ratios of numerals

BARBARA J. ROSE

University of California, Los Angeles, Los Angeles, California 90024

and

MICHAEL H. BIRNBAUM

University of Illinois, Urbana-Champaign, Champaign, Illinois 61820

Each subject performed two tasks, dividing a line segment so that either (a) the *ratio* of subjective lengths corresponded to the ratio of the magnitudes of two numerals or (b) the *difference* in length was proportional to the numerical difference. Had subjects actually performed two operations on the same scale, the responses would have been nonmonotonically related. Instead, data for the two tasks were nearly identical and ordinally compatible with either a ratio or a subtractive model. The ratio model implied scale values for numerals that were a positively accelerated function of numerical value, inconsistent with previous results. With a nonlinear response function for graphic length, the subtractive model fit well, yielding scale values that were a negatively accelerated function of numerical value and a linear function of previously obtained scales. These results, together with other recent findings, suggest that subjects may perform the same operation in spite of instructions to judge "ratios" or "differences" and that this operation can be best represented by a subtractive model.

In these experiments, subjects were presented with pairs of numerals and asked to judge both the *difference* in magnitude between the two numerals and the *ratio*.

For the *difference task*, the subjects were instructed to divide a 100-mm line such that the difference in length between the two segments would be proportional to the "difference" between the numerals. The "differences" were initially assumed to follow a *subtractive model*:

$$R_{ij}^D = J(s_i - s_j), \quad (1)$$

where s_i and s_j are the scale values of the i^{th} and j^{th} levels of the stimuli, R_{ij}^D is the difference in graphic rating length (the dependent variable), and J is a monotone function relating subjective differences to graphic ratings.

For the *ratio task*, the subjects were instructed to divide the 100-mm line so that the ratio of the lengths of the two segments corresponded to the "ratio" of numerical magnitudes. The "ratios" were initially assumed to follow a *ratio model*:

$$R_{ij}^R = J^*(s_i/s_j), \quad (2)$$

where s_i and s_j are scale values as above, R_{ij}^R is the ratio of graphic response lengths, and J^* is a monotone function relating subjective ratios to graphic length ratios.

A study that uses only one task (*difference or ratio*) cannot discriminate Equation 1 from Equation 2, since a monotonic (logarithmic) transformation of a ratio yields a difference. However, the simultaneous evaluation of two or more judgmental tasks with the assumption that scale values are independent of the integration function limits the possible data transformations so that it may be possible in principle to reject at least one of the integration functions under consideration (Birnbaum, 1974a; Birnbaum & Veit, 1974). On the positive side, since subjective values derived from Equation 1 are unique to a linear transformation, and values derived from Equation 2 are unique to a power transformation, if a single scale satisfies both equations, it will be unique to a similarity transformation (Krantz, Luce, Suppes, & Tversky, 1971). It is assumed that the values for s derived from Equations 1 and 2 should not only be consistent with one another but should also be a negatively accelerated function of numerical value in agreement with the results of previous research (e.g., Birnbaum, 1974b; Rule & Curtis, 1973).

In the Birnbaum (1974b) study, subjects rated numerals ranging from 108 to 992 and presented in varying frequency distributions; the psychophysical function for numerals was derived from the fit of a model of contextual effects based upon range-frequency theory (Parducci & Perrett, 1971). The

Computing support was provided by Campus Computing Network, University of California, Los Angeles. The second author received support from a National Institute of Mental Health postdoctoral fellowship at the Center for Human Information Processing, University of California, San Diego, which provided additional support through NIMH Grant MH 15828. We thank Norman H. Anderson for criticisms of earlier drafts. Requests for reprints may be sent to either author: Barbara J. Rose, Department of Psychology, University of California, Los Angeles, Los Angeles, California 90024; Michael H. Birnbaum, Department of Psychology, University of Illinois, Urbana-Champaign, Champaign, Illinois 61820.

model describes the ratings as a composition of a judgment function that depends on contextual effects and a psychophysical function assumed to be independent of context. The resulting psychophysical function for numerals was negatively accelerated and was demonstrated to be invariant across contexts. Rule and Curtis (1973) had subjects compare integers from 1 to 9 with the heaviness of weights. Assuming a subtractive model for number-weight comparisons, they derived a function for subjective number. The function for subjective number was approximated as a power function of objective number with an exponent of .63. The Birnbaum (1974b) scale values were nearly linearly related to the Rule and Curtis (1973) function, extrapolated to the larger stimulus range. Rule and Curtis also summarize other reports that have derived similar results for numerals.

EXPERIMENT I

Method

Stimuli. The stimuli were 81 pairs of numerals based upon a 9 by 9 (Left-Side Numeral by Right-Side Numeral) factorial design. The left-side numerals were 114, 263, 391, 409, 538, 686, 774, 853, and 927. The right-side numerals were 196, 257, 326, 433, 579, 612, 761, 820, and 982.

Procedure. The 81 stimulus pairs were printed in random order on five pages to form booklets. Page ordering within booklets used all 120 permutations of the five pages.

For both tasks, a 100-mm horizontal line containing a short vertical mark at the midpoint and at each end point was printed to the right of each stimulus pair. Judgments were made by drawing a vertical mark on the horizontal line. For both tasks, the subjects were instructed not to try to formulate precise rules or to perform numerical calculations. Each subject performed both the difference task and the ratio task using the same set of 81 stimulus pairs.

For the difference task, the printed instructions directed the subjects to divide the 100-mm line so that the difference between the line segments would indicate the difference between the two numerals. The instructions read (in part): "If you judge the number on the left to be larger, then your mark would be to the left of the midpoint. If you judge the number on the right to be larger, then your mark would be to the right of the midpoint. The greater the difference, the further your mark should be from the midpoint."

For the ratio task, the printed instructions directed the subjects to judge the ratio of the numeral on the left to the numeral on the right. "Your mark will divide the line so that the ratio of the length to the left of your mark to the length to the right of your mark will correspond to the ratio of the number on the left to the number on the right."

In the booklet for each task, the subjects were given two general examples and five warm-up pairs: (657 281), (423 704), (135 962), (576 533), and (819 348). The subjects were allowed as much time as they needed to complete each task, requiring between 30 and 50 min for the completion of both tasks.

Subjects. The subjects were 78 University of California, Los Angeles undergraduates fulfilling a requirement in introductory psychology.

Thirty-nine subjects received the difference task first, and 39 subjects received the ratio task first. Since the two tasks required responses in opposite directions for the same numeral pair, it was possible, at least superficially, to check whether the subjects redirected their scales with the change in instructions. Fourteen subjects were eliminated from each task order, most for failing to reverse the scale as instructed, leaving 25 subjects in each group. Analysis of the effect of task order indicated no discernible effect of task order for the data of either task.

Results

Effect of instructions. Figure 1 plots the mean difference in length between the two line segments as a function of the left-side numeral, with a separate curve for each level of the right-side numeral. The similarity of the data for the two panels shows that the instructions to judge "differences" or "ratios" has little effect on the responses. If the subjects were taking both ratios and differences as instructed, Figures 1A and 1B would look quite different from one another and the three-way interaction would be quite large. Instructions by Left by Right interaction was very small, but statistically significant, $F(64,3136) = 1.49$. Considering the power of the test, these data indicate that instructions to judge differences or ratios had little effect.

If only one of these tasks were studied, it would seem reasonable to rescale the data in accord with the model dictated by the instructions; however, the finding that instructions have little effect suggests that both subtractive and ratio models will have to be considered.

All of the following analyses were performed separately for each set of data, and the results were nearly identical. Therefore, the following analyses are reported for the combined data.

Test of subtractive model, assuming J linear. With the assumption that J in Equation 1 is linear, the difference in lengths would be directly proportional to the subjective difference in length. If the subtractive model were valid and J linear, then the curves in Figures 1A and 1B would be parallel. Instead, the two panels show consistent violations of this graphical prediction: the lower curves are positively accelerated, and the upper curves are negatively accelerated. The steepness of the curves in the region between 391 and 409 does not influence the interaction but may reflect a tendency for subjects to place greater relevance on the first digit. The nonparallelism of Figures 1A and 1B is demonstrated statistically by large and significant Left Side by Right Side interactions,

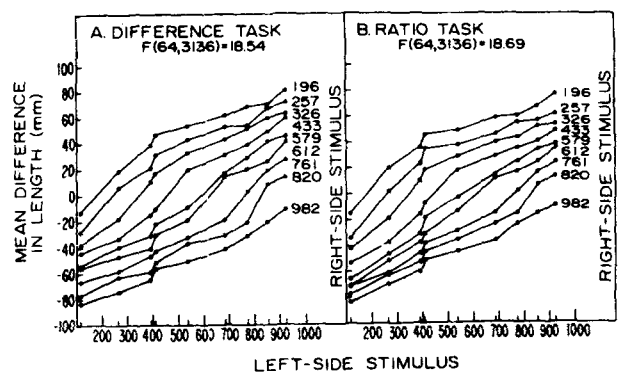


Figure 1. Mean difference in length as a function of the left-side numeral. (A) Difference task data. (B) Ratio task data. (Experiment I.)

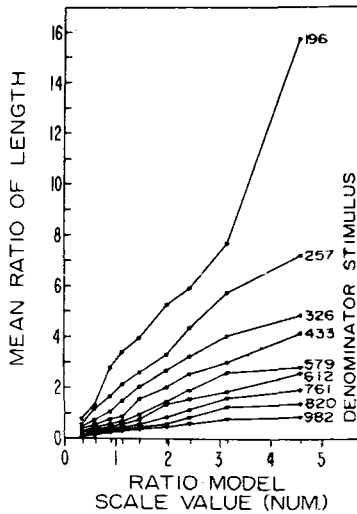


Figure 2. Mean ratios of length as a function of ratio model scale value (marginal mean) for numerator numeral; separate curves represent different numerals for the denominator. (Experiment 1.)

$F(64,3136) = 18.54$ and 18.69 , $p < .001$, for the two respective sets of data. The nonparallelism of the curves in Figures 1A and 1B indicates that either the subtractive model is inappropriate, and/or the assumption of linearity of J is inappropriate.

Test of ratio model. Figure 2 plots the mean ratio of length as a function of the marginal mean for the left-side stimulus (ratio model scale value for numerator number). With the assumption that J^* in Equation 2 is linear, the ratio model predicts a family of linear, diverging curves that intersect at a common point; the interaction should be significant but located entirely in the bilinear component (Anderson, 1970). Consistent with the prediction, the Left by Right interaction is large and highly significant, $F(64,3136) = 22.12$, $p < .001$, with 90.6% and 87.6% of the variance in the bilinear component for the difference and ratio task data, respectively. The residual interactions are small, but statistically significant, $F(63,3087) = 2.61$ and 2.58 , respectively. These residual deviations could easily be removed by monotone transformation as noted below. Even without transformation, the assumption that subjective graphic length is proportional to physical length leads to a reasonably good fit to the ratio model.

The scale values for numerical magnitude, s_i and s_j , derived from the ratio model are plotted in Figure 3 as a function of the physical value of the numerals. The derived function is a positively accelerated function of numerical value, contrary to previous findings of a negatively accelerated function (Birnbbaum, 1974b; Rule & Curtis, 1973). Although the ratio model can be fit to the data, the nonagreement of the scale values derived from the ratio model suggests rejection of this model and/or the assumption of a near-linear

function for graphic rating length (J^* in Equation 2). The fit of the ratio model would be preserved by any power function transformation; the psychophysical function would be changed by the same transformation. It might seem possible, then, to find a power function of graphic rating length that would yield scale values for the ratio model that are compatible with previously obtained results (Birnbbaum, 1974b; Rule & Curtis, 1973). An attempt was made to find such a power function; however, for powers ranging from 1 to .01, the ratio model scale values were a positively accelerated function of the physical value of the numerals and of the previously obtained scales. The following argument shows why this attempt failed. From Equation 2, assuming J is a power function, $R_{ij} = (s_i/s_j)^\alpha$, where α is the power function exponent. Taking logarithms, $\log (s_i/s_j)^\alpha = \alpha(\log s_i - \log s_j)$. Hence, the marginal mean logarithm of the ratio of lengths (marginal mean of Figure 4) should be logarithmically related to the scale values for numerals, irrespective of the value of the exponent. Instead, marginal mean logs were linearly related to the scales previously obtained by Birnbbaum (1974b) and Rule and Curtis (1973). Since power transformations of the ratio model scale values continue to yield positively accelerated psychophysical functions for numerals, acceptance of the ratio model would require a psychophysical function for numerals inconsistent with previous research.

Test of subtractive model, assuming J nonlinear. Figure 4 plots the mean difference in the logarithms of the lengths as a function of the marginal means for the left-side stimulus. If differences in graphic rating length represent larger subjective differences when the response is near the ends of the scale, a logarithmic transformation of the ratio of lengths may rectify the graphic rating responses.

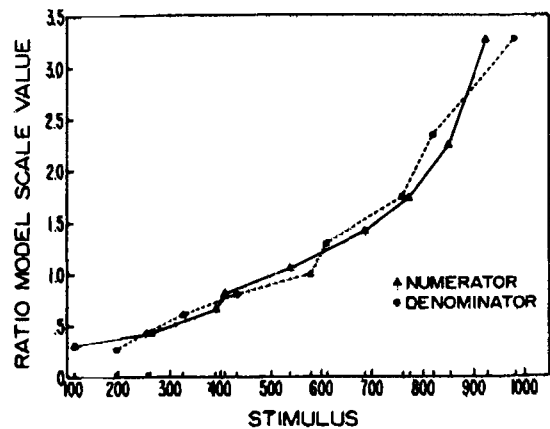


Figure 3. Scale values for the ratio model as a function of the physical values of the numerals; the curve depicts the psychophysical function for numerals derived from the ratio model (Experiment 1.)

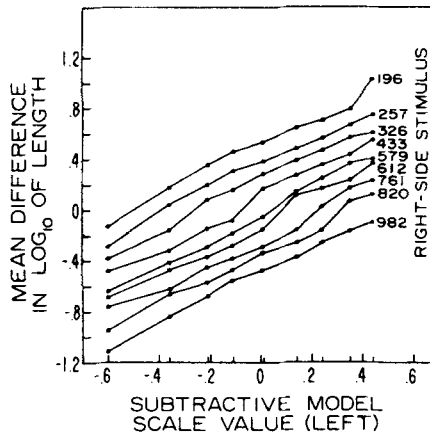


Figure 4. Mean difference in the logarithms of the lengths as a function of the scale value (marginal deviation) for left-side numeral. (Experiment I.)

Following transformation, the curves are approximately parallel, as predicted by the subtractive model. The interaction is small, but very regular and statistically significant, $F(64,3136) = 8.10, p < .01$. The nonparallelism is primarily due to the fact that the curves are steeper in the region of zero on the ordinate; that is, the judged differences are relatively greater for differences of less than 100. The interaction was markedly reduced by Kruskal's (1965) MONANOVA, a computer program for monotone transformation to reduce interactions in analysis of variance (Kruskal & Carmone, 1969). The monotone transformation had little effect on the ordinate of Figure 4, except to reduce the slopes between $-.1$ and $.1$, consistent with the interpretation that the interaction is mainly due to a response tendency to exaggerate small differences. The scale values for the subtractive model (marginal means) were left virtually unchanged by the rescaling. MONANOVA applied to the raw linemark responses yielded equivalent results, since the dependent variables in this constant-sum experiment are monotonically related. Figure 5 plots the rescaled responses, $J^{-1}(R_{ij})$, estimated by MONANOVA. The parallelism of the curves demonstrates that MONANOVA was successful in finding a J transformation to fit Equation 1.

The scale values derived from the subtractive model are plotted in Figure 6 against the physical values. The function is virtually identical to the context-free psychophysical function (dashed curve) derived from range-frequency theory by Birnbaum (1974b). The subtractive model (with the assumption of a nonlinear J function) thus satisfies both the requirement of fit and of scale convergence. Unless one were willing to reject or redefine range-frequency theory, the subtractive representation would be preferred to the ratio representation.

In short, Experiment I suggests that subjects may not distinguish "differences" from "ratios." The assumption of scale convergence leads to the

conclusion that the subtractive representation be retained for both tasks.

EXPERIMENT II

The purpose of the second experiment was to examine the effects of special instructions and training in the numerical and graphical properties of ratios and differences. It was thought that such training might lead the subjects to use two distinct integration strategies with common psychophysical functions. If the special training affected the data for one task, but not for the other, it would seem likely that the subjects had not understood the instructions for that task in Experiment I. The second experiment contained a brief arithmetic test to assess whether subjects understood the distinction between the terms "difference" and "ratio."

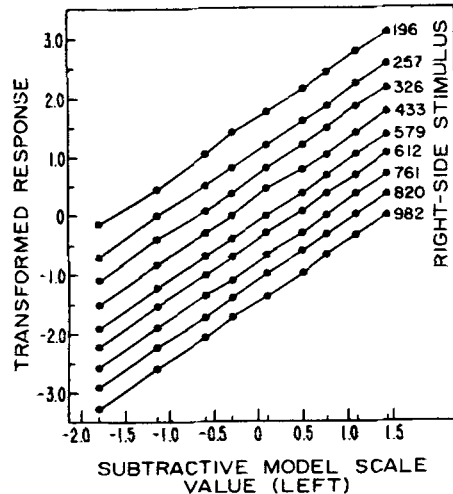


Figure 5. Transformed response, from MONANOVA, as a function of subtractive model scale value. (Experiment I.)

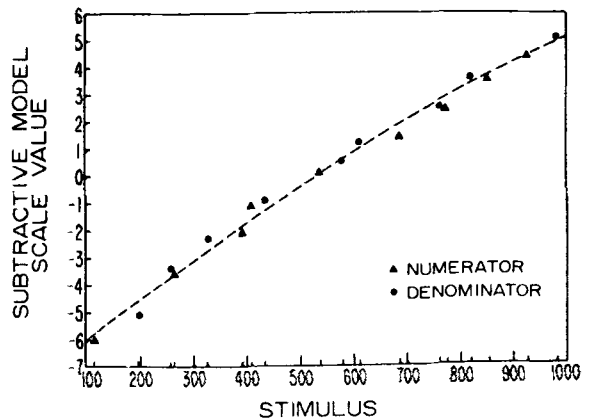


Figure 6. Scale convergence: The plotted points represent the psychophysical function for numerals derived from the subtractive model (Experiment I); the dashed curve is the psychophysical function for numerals derived by Birnbaum (1974b).

Method

Stimuli. The stimulus booklets were identical to those used in Experiment I. Except as noted below, the general procedures were those of Experiment I.

Numerical and graphical training. There were six parts to the training and warm-up that preceded the actual experiment. First, the subjects received a short test in which they were requested to compute ratios and differences numerically. They were given feedback on these items and were given an explanation of simple arithmetic rules. Second, they were given a brief test (without feedback) to ascertain whether they understood the meanings of the terms without additional instruction. Third, they were given a brief lesson in the mathematical and graphical properties of differences and ratios. They were taught, for example, that the difference between 2 and 3 equaled the difference between 3 and 4, but that the ratios were unequal. Subjects were also taught that a ratio of two lengths could be determined by counting the number of times one length fitted into another, and that a difference in length could be determined by cancelling one length from another. This lesson included six examples in which the differences and ratios were worked out arithmetically and graphically. Fourth, there was a brief test to ascertain whether the subjects understood these mathematical properties. Fifth, the subjects read the instructions of Experiment I twice for their first task. Sixth, the subjects completed a set of warm-up trials for their first task.

Procedure. Each subject performed both tasks twice, with half of the subjects performing difference-ratio-difference-ratio, and half in the reverse order. Between each pair of tasks, the subjects read the instructions twice again for the new task and were required to fill out a warm-up. Each change of task required a reversal of the response scale. Each set of warm-ups was checked to determine whether the subject had redirected his scale with the new instructions, and if the subject had not redirected his scale, he was told to reread the instructions and complete another set of warm-ups before proceeding. These procedures resulted in no subject loss.

Subjects. The subjects were 18 University of California, Los Angeles undergraduates who were fulfilling a requirement in introductory psychology.

Results

The mean differences in length are plotted for the two instructional conditions in Figures 7A and 7B, as in Figure 1. The results for the two panels are strikingly similar to each other and to the previous results. The instructions appear to have had little effect, and the three-way Instructions by Left by Right interaction was again of trivial magnitude and was statistically nonsignificant, $F(64,1024) = 1.06$. The Instructions by Left by Right interactions were calculated separately for each subject; $F_s(64,64)$ were less than 1.5 for 13 subjects and less than 2.2 for all 18 subjects. In spite of the special training, subjects do not appear to distinguish "differences" from "ratios."

Consistent with Experiment I, the curves in Figures 7A and 7B demonstrate Left-Side by Right-Side interactions that are highly significant, $F(64,1088) = 24.72$ and 24.77 for the difference task data and the ratio task data, respectively. For each subject, the Left by Right interaction was significant; $F_s(64,64)$ were greater than 1.5 for all subjects and greater than 5.0 for 12 subjects.

All analyses were performed separately for the data of each task with nearly identical results. Every

analysis reported for Experiment I was completed for Experiment II, and the results of Experiment I were replicated in detail. For example, MONANOVA was applied to rescale the difference and ratio task data to parallelism. The rescaled ratings are shown in Figures 8A and 8B, respectively. As in Experiment I, both tasks appear to induce the same ordering, and the scales derived from the subtractive model again agree with the negatively accelerated function for numerals. To fit the ratio model would require exponential transformation, which would yield scale values for numerals that are positively accelerated functions of the previously obtained scales and of objective numerical value.

The remarkable similarity of Figures 1 and 7 and of Figures 5 and 8 shows that both experiments support the same conclusions.

DISCUSSION

In both experiments, subjects appear to perform the same operation regardless of instructions specifying "differences" or "ratios." Had subjects computed both ratios and differences of the same scale values, then the responses would have been

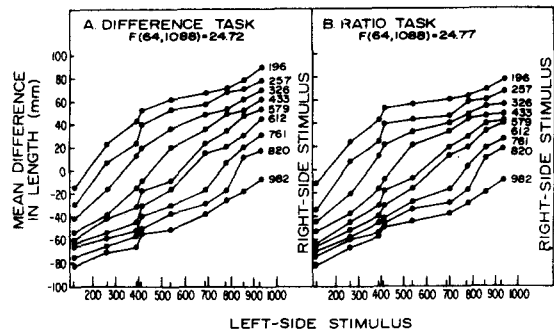


Figure 7. Mean differences in length plotted for the two instructional conditions as in Figure 1. (Experiment II.)

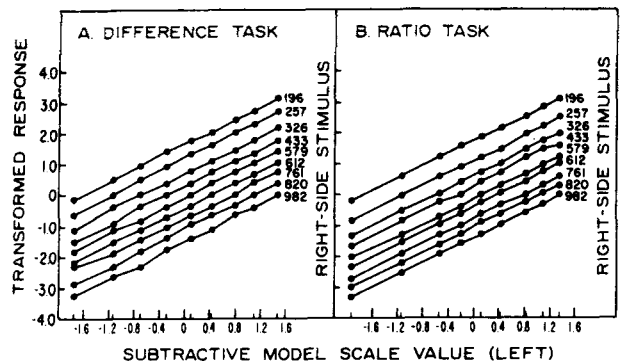


Figure 8. Transformed response, from MONANOVA, as a function of subtractive model scale value for left-side numeral; separate curves represent different right-side numerals. (A) Difference task data. (B) Ratio task data. (Experiment II.)

different for the two tasks. In fact, they would have been nonmonotonically related (Birnbaum & Veit, 1974; Krantz et al., 1971). Instead, Figures 1 and 7 show little difference between the data for the two tasks. Experiment II ascertained that the subjects understood intellectually the distinction between the terms "difference" and "ratio," being able to compute them mathematically. But when asked to perform this task subjectively, their responses appear to be largely unaffected by the task. Since both models cannot be fit to their respective sets of data with a single scale, and there appears to be only one operation, the problem shifts to one of choosing between the models.

In order to assess whether the subtractive or ratio model provided a better representation, two constraints were applied simultaneously: first, the model must be consistent with the data; second, the inferred psychological scale for numerals must agree with the scale derived in previous research. These two constraints led to the conclusion that the subtractive model was appropriate and that the J function for graphic length is nonlinear.

It would seem inappropriate to assume that the J function for graphic ratings would generalize to all other situations. The graphic rating procedure has been widely used in recent psychological experiments, usually with the assumption that responses are an interval measure of the subjective values (e.g., Leon & Anderson, 1974). However, if the J function for graphic rating is nonlinear, interpretation of these studies would require reconsideration. Assuming the subtractive model, the present results would indicate that small differences in response have greater subjective value near the ends of the scale than near the center.

It might be possible to preserve a linear psychophysical function for graphic length by postulating some other form of integration function.¹ However, such complications seem unattractive given the agreement between scales derived from the simple subtractive model and the previous work.

Constant-sum method. In the constant-sum method (Comrey, 1950; Guilford, 1954), subjects are typically instructed to divide 100 points among two or more stimuli to represent the ratios among the stimuli. The scaling of the stimuli is based upon two assumptions: (a) that the ratio model is the appropriate representation and (b) that the subjective magnitude of the number of points is directly proportional to the number. The present results suggest that both assumptions could easily be in error. When analyzing the results of the constant-sum procedure, the possibility that subjects divide subjective values of the points to represent subjective differences should also be considered.

Functional measurement vs. correlational analyses. Goodness-of-fit is frequently assessed in psychophysical research by reducing equations to linear

functions and computing correlations of fit. Often, these analyses are based upon a priori assumptions about the form of the psychophysical functions. Such correlations can be large in spite of serious discrepancies from the model, and inappropriate comparisons of correlations can lead to incorrect conclusions (Birnbaum, 1973, 1974c). For example, in spite of highly significant discrepancies due to the inappropriate assumption that J is linear, the subtractive model achieved a correlation of .99 for Figure 1. The situation is even worse, for the marginal means of Figure 1 showed a near-linear relationship to the previously obtained psychophysical function for numerals. This means that it would be possible to have good scale convergence between situations in spite of serious model discrepancies. Functional measurement (Anderson, 1970) provides more appropriate techniques for testing models. Graphical tests like those in Figures 1, 2, 4, 5, and 8 provide more adequate assessment of the appropriateness of the model than do plots of predicted vs. obtained, which can be misleading (Birnbaum, 1973, 1974c).

Several tasks and scale convergence. The finding that the subjects were apparently unaffected by instructions is consistent with previous conjectures that subjects may not distinguish between "ratios" and "differences" (Birnbaum & Veit, 1974; Torgerson, 1961). Birnbaum and Veit (1974) found that ratings of differences and estimations of ratios of heaviness were monotonically related, inconsistent with the hypothesis that subjects compute ratios and differences on the same scale. Garner (1954) found that subjects made the same settings when instructed to adjust a tone to bisect a loudness interval or to establish equal ratios. Veit (1975) observed that "ratios" and "differences" of grayness of pairs of gray chips were monotonically related. On the basis of a series of experiments, she concluded that the subtractive representation was appropriate and that the ratio model could be rejected. Although the possibility remains that the fault may lie with the instructions or the particular tasks, the special training in Experiment II did not appear to affect the results.

Further research is needed to determine whether there are situations in which simultaneous evaluation of additive and multiplicative processes would yield two distinct models with a single psychophysical function. If it were found that subjects could generate ratios and differences of a common scale, the measurements would be unique to a ratio scale. This finding would fulfill a goal of psychophysics: the development of measurements that interlock psychological laws. On the other hand, it may be that subjects perceive only a single relation between a pair of stimuli for many continua. If so, it may be possible to explain the troubling lack of agreement between so-called "ratio" and "interval" procedures.

The present results provide an example of a

situation where the appropriate model cannot be determined from a priori consideration of the instructions. The use of two tasks not only made possible a test for the effect of instructions, but led to the test of two models. The present data in the hands of two separate investigators, each investigating one of these tasks, most likely would have led to acceptance of two *different* models. If the ratio model were hypothesized, a reasonable fit could be obtained with the assumption of a power function for graphic rating length; if the subtractive model were assumed, the model would be fit with a nonlinear function for graphic length. The investigator utilizing a ratio task and hypothesizing a ratio model would present Figures 2 and 3; the investigator with a difference task and subtractive model might present Figures 4 and 6. The two investigators would be puzzled by the lack of apparent agreement between their findings, when in fact the data are identical—only the hypothesis differs. The use of two tasks does not, in this case, resolve the indeterminacy, but it clarifies the distinction between the task and the model and forces one to confront the indeterminacy problem.

A criterion that goes hand in hand with the evaluation of several tasks is the assumption that psychophysical scales should have greater generality than their role in fitting a single model. This assumption is the basis for deciding between the ratio and the subtractive model. The assumption of the validity of the previously obtained psychophysical function for numerals together with the assumption of scale convergence leads to a preference for the subtractive representation for both tasks.

REFERENCES

- ANDERSON, N. H. Functional measurement and psychophysical judgment. *Psychological Review*, 1970, **77**, 153-170.
- BIRNBAUM, M. H. The devil rides again: Correlation as an index of fit. *Psychological Bulletin*, 1973, **79**, 239-242.
- BIRNBAUM, M. H. The nonadditivity of personality impressions. *Journal of Experimental Psychology Monograph*, 1974, **102**, 543-561. (a)
- BIRNBAUM, M. H. Using contextual effects to derive psychophysical scales. *Perception & Psychophysics*, 1974, **15**, 89-96. (b)
- BIRNBAUM, M. H. Reply to the devil's advocates: Don't confound model testing and measurement. *Psychological Bulletin*, 1974, **81**, 854-859. (c)
- BIRNBAUM, M. H., & VEIT, C. T. Scale convergence as a criterion for rescaling: Information integration with difference, ratio, and averaging tasks. *Perception & Psychophysics*, 1974, **15**, 7-15.
- COMREY, A. L. A proposed method for absolute ratio scaling. *Psychometrika*, 1950, **15**, 317-325.
- GARNER, W. R. A technique and a scale for loudness measurement. *Journal of the Acoustical Society of America*, 1954, **26**, 73-88.
- GULLFORD, J. P. *Psychometric methods* (2nd ed.). New York: McGraw-Hill, 1954.
- KRANTZ, D. H., LUCE, R. D., SUPPES, P., & TVERSKY, A. *Foundations of measurement* (Vol. 1). New York: Academic Press, 1971.
- KRUSKAL, J. B. Analysis of factorial experiments by estimating monotone transformations of the data. *Journal of the Royal Statistical Society (B)*, 1965, **27**, 251-263.
- KRUSKAL, J. B., & CARMONE, F. J. MONANOVA: A FORTRAN-IV program for monotone analysis of variance. *Behavioral Science*, 1969, **14**, 165-166.
- LEON, M., & ANDERSON, N. H. A ratio rule from integration theory applied to inference judgments. *Journal of Experimental Psychology*, 1974, **102**, 27-36.
- PARDUCCI, A., & PERRETT, L. F. Category rating scales: Effects of relative spacing and frequency of stimulus values. *Journal of Experimental Psychology Monograph*, 1971, **89**, 427-452.
- RULE, S. J., & CURTIS, D. W. Conjoint scaling of subjective number and weight. *Journal of Experimental Psychology*, 1973, **97**, 305-309.
- TORGERSON, W. S. Distances and ratios in psychological scaling. *Acta Psychologica*, 1961, **19**, 201-205.
- VEIT, C. T. *Ratio and subtractive processes in psychophysical judgment*. Unpublished doctoral dissertation, University of California, Los Angeles, 1975.

NOTE

1. One possibly attractive speculation is that subjects may be relating each component to the total, i.e., $s_i/(s_i + s_j)$. This formulation is equivalent to the ratio model for this experiment and also yields scales for numerals incompatible with previous research; therefore, this interpretation would also be rejected.

(Received for publication January 13, 1975;
revision received May 12, 1975.)