Measurement and the mental map

MICHAEL H. BIRNBAUM and BARBARA A. MELLERS
University of Illinois at Champaign-Urbana, Champaign, Illinois 61820

In order to test between subtractive and ratio theories of stimulus comparison, judges were asked to estimate "ratios" and "differences" of easterliness and westerliness of U.S. cities. "Difference" judgments fit the subtractive model, and "ratio" judgments fit the ratio model. However, "ratios" and "differences" were monotonically related, contrary to the theory that judges compute both relations on a common scale. Results are consistent with the theory that there is but one operation for both "ratios" and "differences." To assume that the single operation is a ratio requires the complex interpretation that easterliness and westerliness are nonlinearly related. A simpler interpretation is provided by a subtractive theory, in which all four types of judgments are monotonically related to subjective differences on a single cognitive map.

Recent research has shown that when judges are asked to make magnitude estimations of "ratios," the ratio model gives a good approximation to the data. When the same judges are instructed to make category ratings of "differences," the subtractive model closely fits the data.[1] Yet scale values for the same stimuli derived from the two models (applied to data for their respective tasks) do not agree. Data for the two tasks are inconsistent with the theory that subjects are computing both differences and ratios on a single scale of sensation. Instead, the responses for both tasks are monotonically related, consistent with Torgerson's (1961) hypothesis that there is but one comparison operation underlying both types of judgment (Birnbaum, 1978; Birnbaum & Elmasian, 1977; Birnbaum & Veit, 1974; Veit, 1978).

Torgerson (1960) found that magnitude estimations of lightness and darkness of Munsell chips were reciprocally related to each other, whereas category ratings of lightness and darkness were linearly related. Magnitude estimations were an approximately exponential function of category ratings. With the method of single stimuli, as in Torgerson (1960), ratios and differences are necessarily monotonically related, whether the judge uses two operations or one, since x/c is monotonically related to $x - c$ if c is a constant. Ratios and differences of stimulus pairs, in which both stimuli are varied independently are not monotonically related (e.g., $2/1 > 7/5$ but $2 - 1 < 7 - 5$). Consequently, the findings of Birnbaum and Elmasian

(1977) and Birnbaum and Veit (1974) that judgments of "ratios" and "differences" are monotonically related in spite of the fact that the stimulus pairs are constructed from factorial designs provides stronger evidence for the hypothesis that judges use only one operation. They could be computing either differences or ratios—but not both.

The present experiment tests implications of ratio vs. subtractive theories of the comparison operation by having subjects judge "ratios" and "differences" of easterliness and westerliness of U.S. cities. For example, judge the ratio of the easterliness of Philadelphia to that of San Francisco.

But how can one judge ratios of position?

Ratio Theory

One might think that judges would be unable to make sense of the "ratio" task, since ratios require a zero point to be meaningful (Suppes & Zinnes, 1963). However, one strategy for judging "ratios" of easterliness would be to compare mental map distances from the two cities to an arbitrary zero point inserted, perhaps, in the Pacific Ocean. Judgments of "ratios" of easterliness would then be analogous to ratios of line lengths. Similarly, "ratios" of westerliness could be computed by comparing map distances measured from another zero point in the Atlantic Ocean.

This ratio theory implies that judgments of "ratios" will be given by the equations:

$$RE_{ij} = J_{RE}[(s_i - e)/(s_j - e)] \qquad (1)$$

and

$$RW_{ij} = J_{RW}[(w - s_i)/(w - s_j)], \qquad (2)$$

where RE_{ij} and RW_{ij} are judgments of "ratios" of

easterliness and westerliness between cities i and j, with subjective locations s_i and s_j; e is the zero point (presumably in the Pacific) for easterliness judgments, and w is the zero point (presumably in the Atlantic) for westerliness judgments; J_{RE} and J_{RW} are strictly monotonic judgment functions relating overt responses to subjective impressions. If the J functions are linear, it follows that the raw data will satisfy numerical tests of bilinearity and the marginal means, $RE_{i.}$ and $RW_{.j}$ are linearly related to the scale values. Therefore, if "ratios" are computed in this fashion, with a single cognitive map, scale value estimates for "ratios" of easterliness and westerliness should be linearly related, since the scales differ only in origin and the direction of measurement.

Subtractive Theory

Under a subtractive theory, "ratios" of easterliness should pose no particular problem for the judge, since subtraction does not require a zero point to define meaningful intervals. Birnbaum (1978) and Veit (1978) theorized that when the subjective stimulus representation is inherently no more than an interval scale, judges may compare two stimuli by subtraction whether instructed to judge a "difference" or a "ratio." They postulated that degrees of heaviness, loudness, darkness, or likeableness are analogous to locations of points on a line like the longitudes of cities on a cognitive map. To account for the fact that "ratio" judgments are in close numerical agreement with the bilinear form predicted by the ratio model, Birnbaum and Veit (1974) hypothesized that the J function for magnitude estimation is approximately exponential.[2] The subtractive theory predicts that "ratios" of easterliness should fit a ratio model, since an exponential transformation converts subjective differences into numerical ratios.

The subtractive theory is represented in Figure 1. When two stimuli, i and j, are presented for comparison, the judge computes the difference between their scale values, regardless of whether the task is to judge "differences" or "ratios." The response procedure affects the judgment function, J. If the task is to make a *rating* of the "difference," the function is approximately linear. If the task is to make a *magnitude estimation of* "ratio," the judgment function is approximately exponential.

"Ratios" of easterliness and westerliness for the subtractive theory would be given by the equations:

$$RE_{ij} = J^*_{RE}[s_i - s_j] \qquad (3)$$

and

$$RW_{ij} = J^*_{RW}[s_j - s_i], \qquad (4)$$

where RE_{ij}, RW_{ij}, s_i, and s_j are defined as in Equa-

tions 1 and 2. The judgment functions, J^*_{RE} and J^*_{RW}, are assumed to be exponential.[3] It follows that the numerical judgments of "ratios" will fit the ratio model, for $RE_{ij} = \exp(s_i - s_j)$ implies that $RE_{ij} = \exp(s_i)/\exp(s_j) = s^*_i/s^*_j$, where $s^* = \exp(s)$.

The subtractive theory implies that scales estimated from the ratio model (e.g., marginal means of the "ratio" judgments) will be exponentially related to their "true" scale values in the subjective, mental map. If $RE_{ij} = \exp(s_i - s_j)$ and $RW_{ij} = \exp(s_j - s_i)$, then $RE_{i.} = k\exp(s_i)$ and $RW_{i.} = k'\exp(-s_i) = k'/\exp(s_i)$; therefore, $RW_{i.} = kk'/RE_{i.}$, where k and k' are constants. Hence, the subtractive theory predicts that easterliness scale values estimated from the ratio model will be erroneous and nonlinearly (reciprocally) related to the ratio model scales for westerliness. The subtractive theory predicts that scale values cannot be estimated from marginal means of magnitude estimations of "ratios" until the raw judgments are transformed to fit the subtractive model. Thus, the subtractive theory implies that all four tasks—"ratios" and "differences" of easterliness and westerliness—can be represented by differences between cities on a single cognitive map.

This experiment tests between two theories of stimulus comparison. The ratio theory (Equations 1 and 2) implies that scale values derived from the ratio model for easterliness and westerliness should be linearly related. The subtractive theory (Equations 3 and 4), in contrast, predicts that these two ratio model scales will be nonlinearly (reciprocally) related.

METHOD

The subjects performed four tasks, judging "differences" (D) and "ratios" (R) of easterliness and westerliness of cities in the United States. An equal number of judges performed the experiment in each of four task orders, RD-RD or DR-DR by easterliness or westerliness first.

Stimuli and Design

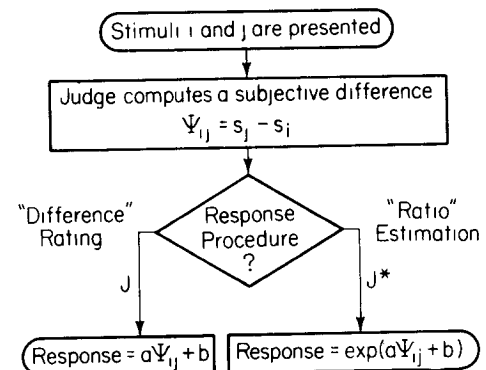The seven cities were San Francisco, California (SF); Salt Lake

Figure 1. Outline of the subtractive model for "ratio" and "difference" tasks.

City, Utah (SLC); Denver, Colorado (Den); Kansas City, Kansas (KC); Champaign-Urbana, Illinois (CU); Columbus, Ohio (Col); and Philadelphia, Pennsylvania (Phil). The pairs of cities were constructed from a 7 by 7, First City by Second City, factorial design, in which the same seven cities were the levels of each factor.

Procedure

The 49 trials were printed in random order in booklets. The first two pages contained instructions, followed by 14 representative warm-up trials.

The "difference" instructions asked the judge to rate the "differences" in the westerliness (or easterliness) between pairs of cities. For westerliness judgments, the subjects were instructed to use integers from 80 (the first city is very very much more westerly than the second city) through 0 (there is no difference in westerliness between the cities) to -80 (the second city is very very much more westerly than the first city. For easterliness judgments, the word "easterly" replaced "westerly" throughout.

The "ratio" task called for estimations of the "ratios" of easterliness (or westerliness) of the cities, using a modulus of 100. Judges were instructed to respond "12.5" if the first city were 1/8th as easterly as the second city; "25" if it seemed 1/4th as easterly; "50" for 1/2; "100" if both cities were equally easterly (e.g., San Francisco: San Francisco); "200" if the first city were twice as easterly; "400" for four times; and "800" for eight

times as easterly. Judges were encouraged to use values between the examples or more extreme ones to represent subjective "ratios" of easterliness (or westerliness).

After completing the four "ratio" and "difference" tasks, the judges were given an outline map of the United States, on which only state lines were indicated. They marked points on the map to represent their beliefs about the cities' locations.

Subjects

The judges were 46 University of Illinois at Urbana-Champaign undergraduates, who received credit in lower division psychology courses. An addition 17 undergraduates participated in a pilot study, using the same stimuli and procedure. Separate analyses of both groups were quite similar; hence, analyses reported here are based on pooled data for 63 subjects.

Eight other judges gave unusual data that were considered separately. Two of these wrote an integer from 1 to 7 over each city's name, calculated ratios and differences numerically, and multiplied by 100 or 10, respectively. The other six gave data that were unreliable or inconsistent with instructions.

RESULTS

Judgments of easterliness are plotted in the upper three panels of Figure 2; judgments of westerliness
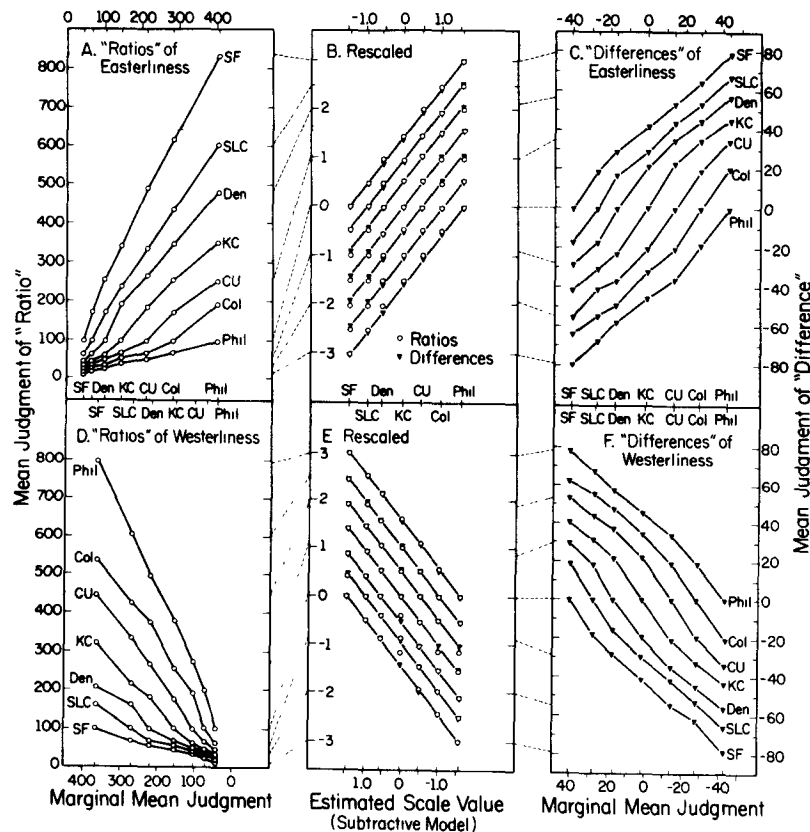

Figure 2. Mean judgments of easterliness and westerliness. "Ratio" judgments (open circles in Panels A and D) are plotted as a function of the marginal means of the first city with a separate curve for each second city. "Difference" judgments (solid triangles in Panels C and F) are plotted in the same fashion. The center panels show data from both tasks rescaled to fit the subtractive model, plotted as a function of the averaged rescaled values for "ratios" and "differences." Note that rescaled values for the two tasks are monotonically related. Lines connecting the panels illustrate the transformations to parallelism.
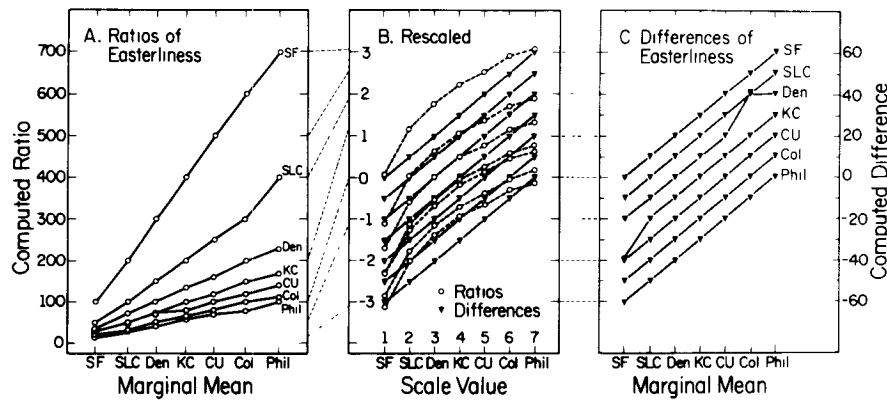
Figure 3. Numerically computed responses from one subject, very similar to the theoretical predictions assuming two operations on a single scale of sensation. "Ratio" judgments in Panel A (plotted as in Figure 2) fit the ratio model, and "difference" judgments fit the subtractive model. However, for this judge, rescaled values from both tasks (center panel) do not coincide; there are two distinct orderings. Only two subjects showed this pattern.

are plotted in the lower panels. "Ratios" are shown in Panels A and D, and "differences" are shown in Panels C and F. The center panels, B and E, show monotonically rescaled means, transformed to fit the subtractive model with a single set of scale values for all four tasks. Lines connecting the panels illustrate the monotonic transformations.

"Ratio" estimations are plotted as a function of the estimated scale values for the first city (marginal means), with a separate curve for each second city. The curves form bilinear fans consistent with the predictions of the ratio model. "Difference" ratings, when plotted in the same fashion, form a set of nearly parallel curves, consistent with predictions made by a subtractive model.

Judgments of easterliness and westerliness could be monotonically transformed to parallelism as shown in Figure 2, Panels B and E. The straight lines are predictions based on the theory that "ratios" and "differences" are both computed by subtraction on a single cognitive map.[4] There are two important points to note: First, the rescaled values for the "difference" data (solid triangles) fall extremely close to the rescaled "ratio" judgments (open circles). The fact that the two sets of orderings for "differences" and "ratios" are nearly identical is consistent with the theory that judges are using the same comparison process for both "ratios" and "differences." Second, the circles and triangles fall very close to predictions (straight lines) derived from the assumption that there is one set of scale values for all four tasks.

If "difference" and "ratio" judgments had been computed numerically, the two operations would have led to two distinct orderings. In this case, it would not have been possible to find monotonic functions that would make the rescaled "differences"

coincide with rescaled "ratios." This possibility is illustrated in Figure 3, which shows the data for one of two subjects who placed an integer from 1 to 7 above each of the cities and attempted to calculate responses. The "ratio calculations" fit the ratio model (Panel A) and the "differences" fit the subtractive model (Panel C), as in Figure 2. However, in this case, data from the two tasks could *not* be transformed to a single ordering (see Panel B). It is interesting that the only two subjects who appeared to have two operations wrote down numbers, apparently to facilitate calculations.

Figure 4 shows the scale values for cities' locations based on ratio and subtractive models. For comparison with the actual map of the United States (Panel D), artist's conceptions have been drawn showing spacing of cities for the ratio model (Panels A and B) and for the subtractive model (Panel C). The maps have been adjusted to match end points.

Panels A and B of Figure 4 show that the ratio model requires different spacings of the cities, depending on the direction of judgment. For example, "ratios" of easterliness imply that Champaign-Urbana (CU) is closer to San Francisco (SF) than to Philadelphia (Phil). "Ratios" of westerliness imply the opposite.

The ratio model yields scale values for easterliness that are reciprocally related to scale values for westerliness, consistent with the interpretation that these comparisons are made by subtraction with an exponential judgment function for magnitude estimation.[5]

Figure 4C shows that the subtractive model is compatible with a single cognitive map. Marginal means for "differences" of easterliness and westerliness are shown as left- and right-facing triangles. The solid circles (to which the arrows point) are based on the rescaled data for all four tasks. These scale values,

A. Ratio Model: Easterliness   C. Subtractive Model: One Map

▷ East
◁ West
• Rescaled

□ Actual
■ Centroid
Response

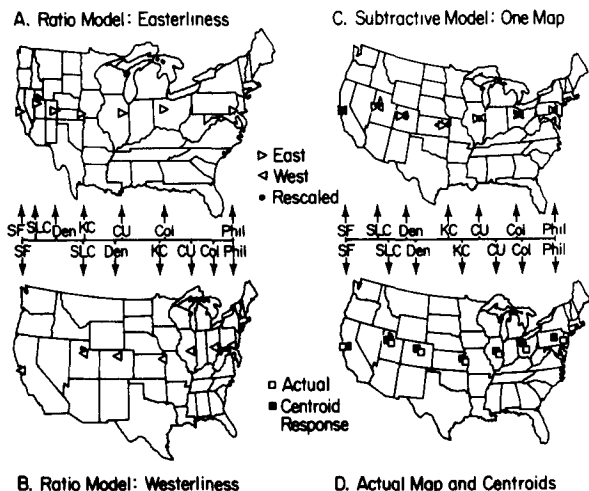B. Ratio Model: Westerliness   D. Actual Map and Centroids

Figure 4. Panels A and B show the two different mental maps for "ratios" of easterliness and westerliness based on ratio model. Panel C plots easterliness and westerliness values for the subtractive model on the same map; solid points are based on all four tasks of Figure 2. Panel D shows an actual map of the United States together with centroid responses for the map-marking procedure. The subtractive model is preferred since it yields a map that is similar to the actual map and is independent of both the task to judge "ratios" or "differences" and of the direction of judgment.

which were used to generate the predicted lines in Figure 3, are similar to the actual locations of the cities (open squares in Figure 4D).

The solid squares in Figure 4D show centroid responses for the procedure in which judges were asked to mark the cities' locations on an unlabeled map. These values are very close to actual map locations, a result that may have been facilitated by presence of the state borders.

Scale values derived from the subtractive model applied to each set of data separately were very nearly equal. Thus, the subtractive model yields a single cognitive map for all four tasks.

DISCUSSION

Consistent with previous research (Birnbaum & Elmasian, 1977; Birnbaum & Veit, 1974), "ratios" and "differences" are monotonically related. A single rank ordering is consistent with the idea that judges are using one operation on one mental map of sensation.

If there is but one operation, is it better represented by a subtractive or a ratio model?

To save the ratio model and the concept of one cognitive map, it would have to be argued that easterliness is the reciprocal of westerliness and that subjects compute ratios when instructed to judge "differences" by reporting a number logarithmically

related to the subjective ratio. It would not be possible to resolve the "original" from the reciprocal or to find the "true" zero point. The scale values for both easterliness and westerliness are nonlinearly related to the actual map (Figure 4, Panels A and B). Which, then, is the "true" subjective map and which the inverse? These complexities make the ratio model implausible.

It seems desirable to represent the subjective locations of the cities on a single map, with invariant scale values rather than scale values that shift nonlinearly for easterliness and westerliness. If judges are using a single cognitive map, the ratio model scale values are erroneous, since they require a stretching and shrinking of the map for judgments of easterliness or westerliness. Subtractive model scale values *are* consistent with the notion of a single cognitive map. When data are rescaled to fit the subtractive model, scale values are independent of task and direction, and they are more closely related to the actual map and to the centroids of the judged locations of cities (Figure 4, Panels C and D).[6]

Birnbaum (1978) and Veit (1978) suggested that when the stimulus representation is inherently no more than an interval scale, judges may compare two stimuli by subtraction whether instructed to judge "differences" or "ratios." To some, it might seem plausible that such dimensions as heaviness and loudness could have subjective zero points, since the physical dimensions of weight and sound pressure do. It might seem less plausible that a dimension such as easterliness would have a zero point. Our present results show no evidence for making such a distinction: data for easterliness are similar to those for loudness (Birnbaum & Elmasian, 1977) and heaviness (Birnbaum & Veit, 1974). Given only the data, a scientist would be hard-pressed to say whether the judgments were of heaviness, loudness, or easterliness.

The comparison of easterliness with westerliness judgments favors the subtractive model, since the subtractive model allows easterliness and westerliness to be linearly related. The finding that "ratios" and "differences" of either easterliness or westerliness give the same pattern of results as judgments of heaviness (Birnbaum & Veit, 1974), loudness (Birnbaum & Elmasian, 1977), lightness-darkness (Veit, 1978), or likeableness (Birnbaum, 1978) suggests that subtraction is the comparison operation for all these continua. Thus, degrees of loudness or heaviness, for example, may be analogous to *positions* of cities on a cognitive map. In this representation, differences are meaningful, but ratios are not.

In sum, there appears to be one operation for "ratios" and "differences" of easterliness and westerliness. This operation seems best represented by the subtractive model, which yields a single cognitive map.

REFERENCES

BIRNBAUM, M H Differences and ratios in psychological measurement In N J. Castellan & F Restle (Eds.), *Cognitive theory* (Vol 3). Hillsdale, N.J: Erlbaum, 1978.

BIRNBAUM, M H , & ELMASIAN, R. Loudness "ratios" and "differences" involve the same psychophysical operation *Perception & Psychophysics*, 1977, **22**, 383-391.

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.

KRUSKAL, J. B., & CARMONE, F. J. MONANOVA: A FORTRAN-IV program for monotone analysis of variance. *Behavioral Science*, 1969, **14**, 165-166.

SUPPES, P., & ZINNES, J. L. Basic measurement theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1). New York: Wiley, 1963.

TORGERSON, W. S. Quantitative judgment scales. In H. Gulliksen & S. Messick (Eds.), *Psychological scaling Theory and applications*. New York: Wiley, 1960.

TORGERSON, W. S Distances and ratios in psychological scaling. *Acta Psychologica*, 1961, **19**, 201-205.

VEIT, C. T. Ratio and subtractive processes in psychological judgment. *Journal of Experimental Psychology: General*, 1978, **107**, 81-107.

NOTES

1. Quotation marks are used throughout to denote instructions to judge "ratios" or "differences" or numbers obtained with such instructions; quotations are not used for ratio and subtractive models or theoretical statements.

2. Birnbaum and Veit (1974) argued that if subjects compare stimuli by subtraction, the magnitude estimation instructions can induce an exponential transformation from subjective value to numerical response. Suppose a certain stimulus pair (a,b) seems a "large" subjective difference, and the judge calls it "four times." If the same pair is presented in reverse order (b,a), the judge is compelled to call it "one fourth." However, the numerical responses of ¼ and 4 are not equidistant from 1 ("no difference") even though the subjective differences are equal. Thus, if the judge uses a geometrically spaced set of numbers for subjectively equal-spaced categories, an exponential transformation will be induced by the magnitude estimation procedure

3 It is only necessary to assume that the J functions in Equations 1, 2, 3, and 4 are monotonic It is possible to use monotone rescaling to solve for scale values which reproduce the rank order of the data. Computer programs such as MONANOVA (Kruskal & Carmone, 1969) can be used to find the rescalings, interpreted as the inverses of the J functions, and to solve for scale values. The theoretical implications discussed in the text (for the case where J is known) apply to scale values based on a purely ordinal analysis as well.

4. The four sets of data were separately transformed to parallelism via MONANOVA. The eight estimates (row and column by ratio and difference by easterliness and westerliness) for each scale value were highly consistent. The eight estimates were averaged to yield the single set of scale values used in Figures 2B and 2E. Final transformations shown in Figure 2 were accomplished by using a graphical method.

5. For the two judges who calculated with numbers, however, scale values for "ratios" of easterliness and westerliness were linearly related to each other and to the "difference" scale values (see Equations 1 and 2 and abscissa of Figure 3).

6. The theory that judges compute differences followed by an exponential response transformation is mathematically equivalent to the theory that they exponentiate, then compute ratios. The latter interpretation seems complicated, however, since it postulates that the subjects use one mental map to generate two others, depending on the direction of judgment. It seems simpler to assume one map, one comparison operation, and two judgment functions than to assume two operations, and three mental maps related by an intervening exponential transformation. Furthermore, the theory that the judgment function for magnitude estimation is positively accelerated correctly predicts that magnitude estimations of "differences" and "averages" show a divergent interaction that requires rescaling. To argue that magnitude estimation induces an exponential transformation in the stimulus scales for these experiments would also require the interpretation that the comparison and combination processes are dependent upon the procedure for responding (Birnbaum, 1978; Veit, 1978).