Communication & Science
JOURNAL
Classic Series 2012

Sandhaus, D. A., Fink, E. L., Kaplowitz, S. A. (1987). Asymmetry in judgments of similarity.

A RAH Press Publication
http://www.comSciJ.com

Asymmetry in Judgments of Similarity

Douglas A. Sandhaus Edward L. Fink

Department of Communication Arts and Theatre

University of Maryland

College Park, MD 20742

(301) 454-5122

Stan A. Kaplowitz

Department of Sociology

Michigan State University

East Lansing, MI 48824

(517) 355-6634

May 1, 1987

Running Head: ASYMMETRY IN JUDGMENTS OF SIMILARITY

We would like to thank Vicki Freimuth, Theresa Marron, Kenneth Mihalek, and Charles Munro for comments on and/or assistance with this study.

Asymmetry in Judgments of Similarity

Abstract

Asymmetry in judgments of similarity means that "a is similar to b" is perceived differently than "b is similar to a." Tversky and Gati (1978) hypothesize that asymmetry results when the judgment task is directional ("how similar is a to b" is directional, while "how similar are a and b" is nondirectional) and when the concepts judged differ in prominence. A multidimensional model, however, posits that similarity judgments may be seen as distance estimates, and hence, should be symmetric. An experiment was conducted to assess Tversky & Gati's model of asymmetry judgments, using a nondirectional task. 287 subjects participated in a 2 (whether subjects estimated similarity or difference) x 2 (using either a magnitude or a closed-ended scale) x 2 (whether features were made salient before or after subjects made their estimates) x 2 (whether each of twenty-one pairs of concepts were presented with the more prominent concept as the subject or referent of the comparison) design. Overall, we find no support for significant asymmetries in the data. Further, Tversky and Gati's feature contrast model for similarity and difference judgments was not supported.

> "If the distance from Helm to Shedlitz is four miles, what is the distance from Shedlitz back to Helm?"
>
> "Eight miles," came Gimpel's prompt reply.
>
> "And can you tell us why?" asked Mottel the Mayor, smiling eagerly.
>
> "Certainly," Gimpel answered. "It's simple. There are four months from Chanukah to Passover and eight months from Passover back again to Chanukah." (Simon, 1955, p. 5)

Judgments of similarity are at the center of human mental processing. Individuals classify objects, form concepts, make generalizations, and memorize and retrieve information, based on similarity judgments (Tversky, 1977). Ortony (1979) notes that both metaphorical language and metaphorical thought are based on statements and judgments of similarity (p. 162), and if, as Lakoff and Johnson (1980) assert, metaphor may hold the "key to giving an adequate account of understanding" (p. ix), questions about judgments of similarity may be construed as issues about the nature of understanding. To construct a model of similarity relations may be to construct a model of thought itself.

Tversky and Gati (1978) note that the theoretical analysis of similarity relations has been dominated by geometric models. As

Marron (1985) indicates, geometric models have been used to study language, campaign communication, implicit theories of personality, person perception, stereotypes, the structure of emotions, and the structure of society. Geometric models "represent objects as points in some coordinate space such that the observed dissimilarities between objects correspond to the metric distances between the respective points" (Tversky, 1977, p. 327). Most of these models proceed from three assumptions:

(1) The assumption of minimality; that is, that the distance from a point (object) to itself is zero.

(2) The assumption of symmetry; that is, the distance from some point a to a second point b is equal to the distance from b to a.

(3) The assumption of the triangle inequality; that is, the distance from a to b, added to the distance from b to some third point c, will be greater than or equal to the distance from a to c.(1)

In 1977, Tversky formulated his feature-theoretic model as a response to the geometric model. The bulk of Tversky's work, particularly Tversky (1977) and Tversky and Gati (1978), seems to directly contradict the assumption of symmetry. By challenging the assumption of symmetry the implication is that judgments of similarity are not consistent with a geometric model.

Tversky claims the feature-theoretic model can account for the failure of the geometric model to meet the assumption of

symmetry. The feature-theoretic model proposes that concepts are not mere points (or, for that matter, regions) in space, but rather, they are made up of sets of features. The similarity of two concepts a and b is a function their common and distinctive features. Similarity is expressed as a mathematical combination of three components:

(1) $A \cap B$: the features shared by a and b;

(2) $A - B$: the features possessed by a, but not by b;

(3) $B - A$: the features possessed by b, but not by a;

where lower case letters refer to the concept, and upper case letters refer to the features associated with the concept. Tversky and Gati (1978) predicted that judgments of similarity are based upon measures of the degree to which two sets of features match each other, rather than on the metric distance between points in a coordinate space. The similarity of a to b is described as "a linear combination (or a contrast) of the measures of their common and distinctive features" (p. 80). The above components, taken together and labeled the "contrast model," yield a similarity scale S which is written

$$S(a,b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A) \qquad (1)$$

where θ, α, β, the parameters reflecting the salience or prominence of the various features, are all non-negative.

Thus, $S(a,b)$ is expected to increase as the number of similar features is increased [$f(A \cap B)$], and is expected to decrease as the number of distinctive features is increased. When one is

asked to estimate how similar *a* is to *b*, Tversky and Gati (1978) note, "one naturally focuses on the subject of the comparison," i.e., *a* (p. 85). Thus, they assume that the features of the subject are weighted more heavily than the features of the referent (i.e., $\alpha > \beta$). Tversky and Gati demonstrated that the contrast model predicts symmetry ($S(a,b) = S(b,a)$ only when $\alpha = \beta$, or $f(A - B) = f(B - A)$.

In developing their contrast model, Tversky and Gati (1978) presented subjects with twenty-one pairs of countries. Each pair contained a significantly and comparatively more prominent country *p*, and a less prominent country *q*. Subjects were asked to make a directional comparison, that is, they were asked to indicate how different (or similar) one concept is to another, rather than indicating how different (similar) the two concepts are from (to) each other. Half the subjects in one condition (similarity) were then asked to estimate, on a twenty-point scale, how similar *p* is to *q*; the second half of subjects in the similarity condition were asked to estimate how similar *q* is to *p*. Using the same twenty-point scale, subjects in a second condition (difference, abbreviated here as *d*) were asked to estimate how different *p* is from *q*, or how different *q* is from *p*. Tversky and Gati predicted that $s(p,q) > s(q,p)$, and $d(p,q) < d(q,p)$. Their results generally supported their prediction.

However, in testing their predictions regarding similarity, six of the twenty-one pairs (28%) tested by Tversky and Gati

failed to demonstrate asymmetry in the direction they hypothesized; in testing their predictions regarding difference, three pairs -- or 14% -- so failed. Tversky and Gati made no attempt to explain these failures. One might therefore begin to question the generality of the feature-theoretic approach, and refocus one's attention on the appropriateness of Tversky and Gati's feature-theoretic approach.

In Tversky and Gati's study, subjects were asked to assess either "the degree to which Country *a* is similar to Country *b*" or "the degree to which Country *a* is different from Country *b*." From a geometric perspective, "similar to" and "different from" may be considered asking for distance. Yet, Tversky and Gati's twenty-point, closed-ended scale is not consistent with the way people typically consider distance. Distance is usually measured on an open-ended, zero-to-infinity scale, not a closed-ended, one-to-twenty scale. It is quite possible, therefore, that Tversky and Gati's results were caused by the use of an inappropriate measure of similarity. To see if this is the case, we are replicating their study using both magnitude scales, which are appropriate for measuring distance, and their scales. Doing so will tell us if the asymmetries they found are an artifact of their measures.

A key premise of Tversky and Gati's feature-theoretic reasoning is their contrast model (see Eq. 1). Tversky and Gati present no direct evidence to assess its plausibilty. While they claim that the contrast model is consistent with observed

asymmetries, they have never directly tested it by measuring the common and distinctive features of any concepts. We will replicate Tversky and Gati's study, asking subjects to estimate the similarity (or difference) between concepts. In this way we can directly examine any asymmetries that result. Second, we will estimate the parameters of the contrast model (Eq. 1), to see if features enter into judgments of similarity and difference as Tversky suggests.

Method

Overview

Subjects were provided with one of sixteen different forms corresponding to the sixteen different experimental conditions to be described below. Prior to the experiment, subjects were given a short presentation by an experimenter which briefly explained how to use the various measurement scales in the questionnaires. Subjects were given the remainder of the fifty minute class period to complete the forms.

Four independent variables, each with two levels, were manipulated: (1) whether subjects estimated *similarity* or *difference*, (2) whether subjects were asked about features *before* or *after* they estimated similarity or difference, (3) whether subjects utilized a closed-ended *twenty-point scale* or a *magnitude scale* to estimate similarity or difference, and (4) whether the more prominent country was used as the *subject* or *referent* of the comparison.

In choosing an appropriate design, one issue, that of replication, requires mentioning. The present study may be considered, in many ways, a replication of the work done in Israel by Tversky (1977) and Tversky and Gati (1978). The pilot studies mirror the procedures used by Tversky and Gati. The dependent variable in the present study is each subject's estimates of either the similarity or difference of the same twenty-one pairs of countries used by Tversky and Gati (1978). Also, in the present study, one half of the subjects in each condition were presented with Tversky and Gati's closed-ended, twenty-point scale.

Operationalizing Shared and Distinctive Features

The present study will use Tversky and Gati's contrast model to test Tversky (1977) and Tversky and Gati's (1978) hypothesis that estimates of similarity and difference are based on shared and distinct features. Since Tversky and Gati's contrast model has never been tested empirically, there exists no precedent for operationalizing the concepts of shared and distinctive features. While it is presumed that one may operationalize the shared and distinct features of various concepts in many ways, we operationalized them in what we believed was the simplest way. For each of two pairs of countries, subjects were asked to complete three lists. Those lists were: (1) "[Country a] and [Country b] share the following features," (2) "[Country a] has these features which [Country b] does not have," and (3) "[Country

b] has these features which [Country a] does not have." The number of features listed in each list was considered to operationalize, respectively, the shared features of Country a and Country b ($A \cap B$), the distinct features of Country a ($A - B$), and the distinct features of Country b ($B - A$).

Pilot Studies

Two pilot studies were necessary prior to conducting the present research: (1) one to assess the relative prominence of the pairs of countries, and (2) a pilot study to find an appropriate "yardstick" for the magnitude scale judgments of similarity and difference.

The first pilot study presented to subjects the twenty-one pairs of countries used by Tversky and Gati (1978). Subjects were asked to indicate for each pair the country they preferred to use as the referent when making comparison judgments. Tversky regards this as revealing which member of the pair is more prominent. Table 1 presents the outcome of this pilot study. (In all that follows, the more prominent country in a pair will be italicized). For subsequent analyses of each pair of countries, the country listed under column p in Table 1 was considered more prominent than the country listed under column q. In each case, the country that we found to be the more prominent member of the pair is the same one found by Tversky and Gati to be the more prominent.

Insert Table 1 about here ← See pp. 30-31

The second pilot study was used to construct a suitable yardstick for the present study's magnitude scaling conditions. Subjects were presented with a preliminary yardstick, the pair red-white, which subjects were told represented 100 units of similarity (or difference). Subjects were then asked to estimate on a magnitude scale how similar (or different) forty-five pairs of countries were, compared to the standard difference that they were given (red and white = 100). We wished to select a pair of countries whose degree of similarity and difference were perceived as moderate, and for which there was high agreement as to the level of similarity and difference.

Based on these criteria, the pair Jordan-Philippines was selected as the yardstick pair for the final experiment.

Manipulation of Independent Variables

The four experimentally manipulated independent variables are described below:

Subjects in the similarity condition were asked to estimate how similar the two countries were. Subjects in the difference condition were asked to estimate how different the two countries were. Tversky and Gati point out that such questions may be asked in a directional way (e.g., "how different is a from b?"), or in a nondirectional way (e.g., "how different are a and b from each other?"). The instructions in the present experiment emphasized the nondirectional comparison.

Order of prominence within stimulus pair (i.e., the relative

prominence of the referent as compared to the subject in a given pair of concepts) was manipulated as follows. All twenty-one pairs of countries used by Tversky and Gati (1978) were given to subjects. On all forms, approximately half the pairs of countries were presented with the more prominent country -- as was determined by the first pilot study -- first, and half had the more prominent country second. Across all forms each pair of countries was presented half the time with the more prominent country first, and half the time with the more prominent country last.

Scale was manipulated by giving subjects either a closed-ended twenty-point scale or a magnitude scale to make their estimates of similarity or difference. Half the subjects across all conditions were asked to estimate degrees of similarity or difference using Tversky and Gati's closed-ended twenty-point scale. The other half of subjects made their estimates using magnitude scaling.

The shared and distinctive features of two pairs of countries (England-Ireland and U.S.S.R.-Syria) as perceived by the subjects were measured as indicated above, using a single "features page." Half of the subjects in all conditions were presented with this "features page" before they were asked to estimate any similarities or differences; the other half of the subjects were presented with the features page after they made all these estimates. Asking all subjects to list features allows us to

directly test the contrast model by using a function of the reported numbers of features as the function specified in Eq. 1. Varying whether subjects listed features before or after making similarity judgments can tell us whether the effects expected by Tversky are heightened when features are made salient by asking subjects to list them prior to the judgment task. It is possible that Tversky and Gati's feature model is most likely to be supported when subjects have been consciously thinking about features.

Measurement of Similarity or Difference

The dependent variable, subjects' estimates of similarity or difference, was measured using a magnitude or a closed-ended sale. The scale used in the magnitude scaling conditions had "0" (zero) represent total absence of similarity (or difference), and "100" (one hundred) represent the amount of similarity (or difference) as the given predetermined "yardstick" pair of countries.

The scale used in the closed-ended scaling conditions, the same scale used by Tversky, is a one to twenty scale. Subjects were asked to estimate the same twenty-one pairs of countries used by Tversky on this closed-ended scale where "1" represents little or no similarity (or difference), and "20" represents total or a great deal of similarity (or difference).

Testing the Contrast Model

If a is the subject of a comparison and b is the referent of the comparison, then

$$S(a,b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$$

and

$$S(b,a) = \theta f(B \cap A) - \alpha f(B - A) - \beta f(A - B).$$

Let $\alpha = \beta + K$, and let $S(a;b)$ represent a general distance function for a and b, such that

$$S(a;b) = \theta(f(A \cap B)) - (\beta + K(ORD))(f(A - B)) - (\beta + K(1 - ORD))(f(B - A)),$$

where ORD = 1 if we are assessing $S(a,b)$, and ORD = 0 if we are assessing $S(b,a)$. For the contrast model to be supported, we must find that the above equation for the general distance function is statistically significant in explaining the similarity or difference between two concepts. In addition, if we estimate each of the coefficients in this equation freely (i.e., without constraining to equality those which are predicted to be equal), we should find that the two coefficients designated as K should be equal, and the two coefficients designated as β should be equal. Furthermore, if the pair of concepts we are evaluating shows no asymmetry, then it follows that either $K = 0$ or $f(A - B) = f(B - A)$. To test the contrast model, we create a regression equation which includes the five predictor variables indicated in the equation immediately above. The natural logarithms (after adding 1) of the feature variables are used to create the function to represent feature salience, since we theoretically view the salience of the shared and distinct features of two countries to increase logarithmically with the number of such features

mentioned. In addition, we allow the equation to have an intercept, since the exact metrics for the feature measurements were not fully specified by theory, and the intercept allows for the empirical adjustments in the measurements (e.g., the addition of an arbitrary constant to the number of features prior to taking the logarithm).

Subjects

Subjects were 287 students (138 males, 140 females, 9 subjects didn't supply the information) attending either of two afternoon sections of introductory speech communication at a large state university. Both sections contained approximately the same number of males as females. Researchers frequently use students in this class for research during class time, so the experiment was not viewed as unusual to the participants.

Experimental Procedures

The experiment was conducted during a regularly scheduled class meeting. The experimenter introduced himself as a member of a research staff interested in student perceptions of countries. Prior to distributing the questionnaires, the experimenter gave a short lecture on how to respond to a magnitude scale. The experimenter noted that all subjects would be asked, at some point during the experiment, to employ such a scale, and all subjects were in fact asked for some information on a magnitude scale in the study, regardless of condition. Any subject who might still have had any questions on how to answer the questionnaire was

advised to consult with one of the three assistants who were available.

The questionnaires, whose different forms had been previously placed in a random order, were handed out to the subjects by the three assistants. Once each subject had a form, the experimenter read aloud the instructions on page one. The experimenter stressed that there were written instructions and examples prior to each part of the questionnaire, and asked the subjects to read them carefully before answering the questions. Subjects were told that they had the remainder of the class period, approximately forty minutes, to complete the questionnaire. Most subjects completed the forms in twenty-five to thirty minutes.

Results

Treatment of Missing Data

While missing data encountered for subjects' estimates of similarity and difference were routinely eliminated from further analysis, the definition of missing data for subjects' lists of shared and distinctive features requires some elaboration. On the features page, subjects were given two pairs of countries. For each pair of countries, subjects were asked to provide three lists: (1) the shared features of Country a and Country b, (2) the distinctive features of Country a, and (3) the distinctive features of Country b.

For purposes of defining missing data, each pair of countries was considered separately. For any given subject, for either

given pair of countries, feature data was considered missing if that subject listed neither shared features nor distinctive features.

Preliminary Analyses

First, the data were trimmed: scores greater than 99,998 on any magnitude scale were converted to 99,998. Next, the magnitude scaling data was transformed to correct for the positive skew typically found in such data. The data were transformed by adding 1, and then taking the natural logarithm. Such a transformation tends to make the data homoscedastic.

Reliability

Reliability of the means of the raw scores (or, in the magnitude scaling conditions, the means of the transformed scores) for each of the twenty-one pairs of countries was assessed using Pearson's correlation coefficients in the following way.

We have means on 21 countries in the following eight conditions: similarity-sub-close, similarity-sub-mag, similarity-ref-close, similarity-ref-mag, difference-sub-close, difference-sub-mag, difference-ref-close, difference-ref-mag, where "sub" means that the more prominent country was used as the subject, "ref" means that the more prominent country was used as the referent, "close" refers to use of the close-ended 1 to 20 scale, and "mag" refers to the use of the magnitude scale. These means were correlated across conditions (with an N of 21, the number of pairs of countries) resulting in twenty-eight independent

correlation coefficients.

If data are highly reliable, we would expect a high positive correlation both when the comparison was between two similarity conditions, and when the comparison was between two difference conditions (indicating that subjects generally perceived the measures of similarity or difference in roughly the same way). Similarly, if data are highly reliable, we would expect a high negative correlation whenever a similarity condition was compared to a difference condition (indicating that measures of similarity and difference are basically opposites). This was, in fact, exactly what was found. For the similarity conditions, the six correlation coefficients ranged from .8793 to .9704. For the difference conditions, the six correlations ranged from .8637 to .9332. Finally, for the sixteen correlations comparing a similarity condition to a difference condition yielded correlations ranging from -.7532 to -.9167. All 28 correlations are significant ($p < .001$) and in the correct direction. In short, all the data seems to be highly reliable. The use of the means to assess reliability is appropriate since the test of the asymmetry hypothesis will employ the means.

Asymmetry and Feature Salience

Analyses of variance were used to evaluate asymmetry in the judgments. The mean judgments appear in two tables. Table 2 presents the results from use of a closed-ended scale to estimate similarity and difference. Table 3 presents the results from use

of a magnitude scale to estimate similarity and difference.(2)

Insert Tables 2 and 3 about here See pp. 32-35

Using a closed-ended scale, the symmetry hypothesis was rejected for three pairs of countries, all in the similarity condition: Philippines-Japan, Ceylon-India, and U.S.A.-France. Subjects rated first two pairs significantly less similar when the more prominent country was used as the subject of the comparison than when the prominent country was used as the referent. However, subjects rated the U.S.A.-France pair as significantly more similar when the more prominent country was used as the subject then when it was used as the referent.

Using a magnitude scale, the symmetry hypothesis was rejected in only three pairs of countries, all in the difference condition: China-Albania, England-Ireland and China-N. Korea. Subjects rated all three pairs as significantly more different when the more prominent country was used as the subject of the comparison than when the more prominent country was used as the referent.

Additionally, four correlated t-tests (two in magnitude scaling conditions, one for similarity and one for difference, and two in the closed-ended scaling conditions, again one for similarity and one for difference) were conducted on the means of the twenty-one pairs. The null hypothesis for these tests are that the similarity or difference of the pairs of countries is the same, regardless of which country is listed first in the

comparison. None of the four t-tests, which are reported below in Table 4, proved statistically significant.

Insert Table 4 about here See p. 36

The Tversky Equations

We directly tested the contrast model equation using eight regressions. Using the two pairs of countries for which data on the shared and distinctive features were collected (England-Ireland and the U.S.S.R.-Syria), we adapted Tversky's "contrast model" equation as described above. The equation tests whether the number of shared and distinctive features reported by the respondents affects their judgments of similarity or difference, and whether the placement of the more prominent country in the pair affects the coefficients of Tversky's model. When magnitude scales were used, we took the logarithm of the appropriate similarity or difference measure, since magnitude measures result in residuals which are heteroscedastic, while transforming them allows the data to meet the homoscedasticity assumption necessary for statistical analysis.

None of the four England-Ireland equations was found to be significant, and only in two cases was a single variable found to be a significant predictor. The coefficients for this variable (the number of shared features) support the idea that, when using the magnitude scale, the greater the shared features of England and Ireland, the more similar these two countries were perceived

to be.

None of the U.S.S.R.-Syria equations were found to be significant predictors of the dependent variable; that is, regardless of the scale subjects used to estimate either similarity or difference, the equation utilizing the five predictor variables failed to significantly predict those estimates. In addition, no coefficient was statistically significant.

If we examine the coefficients predicted to be equal by the contrast model, we have mixed results. The coefficients for $f(R - S)$ and $f(S - R)$ in Table 5 should equal $-\beta$, and the coefficients for (order)*$f(R - S)$ and (1 - order)*$f(S - R)$ should equal $-K$. Tables 5 & 6 are on p. 37 Similarly, the coefficients for $f(I - E)$ and $f(E - I)$ in Table 6 should equal $-\beta$, while the coefficients for (1 - order)*$f(E - I)$ and (order)*$f(I - E)$ should equal $-K$. Given the large standard errors for both β and K, we cannot reject the null hypothesis that the alternative estimates for β and K in each equation differ.

Tversky and Gati (1978, p. 80) state that α, β and θ are all greater than zero. If we look at the values only for β, we find that 6 of the 16 estimated values in Tables 5 and 6 are negative. Similarly, α (= β + K), is negative in 12 of the 16 comparisons.

In addition, if α does not differ from β, then, according to Tversky and Gati (1978), the only way asymmetry could be

achieved is if $f(A - B)$ differs significantly from $f(B - A)$. Since there was a significant asymmetry for the England-Ireland pair in the magnitude scale-difference condition, and since neither estimated value for k was significantly different from zero (which means that α and β do not differ significantly), we will examine $f(A - B)$ and $f(B - A)$ for this condition. The mean for the number of distinct features for Ireland (after transformation) is .76 (sd = .55), while the mean number of distinctive features for England is .81 (sd = .56). These two mean values do not differ significantly (t = -1.00, df = 67, p = .32).

Discussion

Replication of Tversky's Findings

We expected that asking for similarity or difference in a nondirectional manner would result in fewer asymmetries than was found by Tversky and Gati (1978). Indeed, while we replicated their basic study except for our instructions, we failed to replicate their findings of significant asymmetries overall (see Table 4). Utilizing the close-ended scale employed by Tversky and Gati, neither the similarity nor the difference condition was found to possess a general asymmetry, either in the direction predicted by Tversky and Gati, or in the opposite direction. Symmetry was also generally found using a magnitude scale in both the similarity and difference conditions.

It should be noted that whether the features were measured

first or last had no major effect on the likelihood of producing asymmetries in the judgments. Similarly, contrary to our suspicion, the magnitude scale was not significantly better in producing symmetries in the judgments.

If we follow the logic in the Tversky and Gati discussion, we should not be surprised that our instructions resulted in symmetric judgments of similarity. However, it is not so clear that these results stem from the feature contrast model that they propose, since the regression equations were not significant, and the coefficients did not strongly support the model. Other interpretations for the symmetries found here need to be examined. One possible difference between the Tversky and Gati study and ours is in the subjects used. Their Israeli subjects considered the prominence of the twenty-one pairs of countries as significantly more polarized than the American subjects used in our pilot study. Across the twenty-one pairs of countries, their Israeli subjects agreed on the more prominent country an average of 92.1% of the time, compared to our American subjects who agreed on the more prominent country only 77.1% (the mean of the percentages in Table 1) of the time.(3) These percentages across the twenty-one pairs of countries are significantly different (t =25.51, df =20, p <.001).

Additionally, on a pair-by-pair analysis of the countries considered more prominent, our American subjects were closer to consensus than Tversky and Gati's Israeli subjects for only two

pairs of countries, U.S.A.-France (92.3% of our subjects agreed that the U.S.A. was more prominent, compared to only 86.8% of Tversky and Gati's subjects) and U.S.S.R.-Cuba (96.2% of our subjects agreed that U.S.S.R. was more prominent, compared to 91.1% for Tversky and Gati's subjects). We note that, in the present study, for the estimates of similarity, judgments concerning the U.S.A.-France distance were significantly asymmetrical, while the U.S.S.R.-Cuba distance was almost significantly asymmetrical ($p = .055$). This may suggest that the more prominence is polarized, the more asymmetrical the judgments.

To test this hypothesis we computed the Pearson product-moment correlation on the twenty-one pairs of countries which compared, by pair, the percentages that subjects chose one country as the more prominent with the asymmetries. For our study, this resulted in four correlations, of which none were significant. For Tversky and Gati's study, this resulted in two correlations, neither of which was significant. Thus, there is no evidence that the more prominence is polarized, the more asymmetrical the judgments.

Testing the Featural Model

Tversky's assumptions regarding shared and distinctive features were tested by eight regression analyses. None of them were found to be statistically significant. However, one variable, in the England-Ireland regression, was found to be a significant predictor of the relevant dependent variable in the

magnitude scale conditions; this variable was the number of shared features of these two countries.

If, as Tversky and Gati (1978) assert, similarity (or difference) is a function of the shared and distinctive features of two concepts, we might well expect all of our regressions to be significant, but we did not. Further, the pattern of the coefficients derived from the feature contrast model do not provide strong support to this model. Why might this be the case?

Perhaps the weakness in the present study is in our operationalization of the shared and distinctive features in testing Tversky and Gati's (1978) contrast model. Neither Tversky (1977) nor Tversky and Gati (1978) ever operationalized their measures of the common and distinctive features of concepts. While we assumed that a simple and reasonable interpretation of a "linear measure" of shared and distinctive features would be a function of the number of such features, it must be noted that we had very little guidance in this matter. Since Tversky and Gati suggest that they wish the function employed to reflect the salience of the features, we believe that the operationalization employed here is reasonable.

Given the above discussion, the most obvious reason that the feature contrast model fails the multiplicity of tests we have subjected it to is because the model is incorrect.

Conclusion

The present study tested Tversky and Gati's assumptions that

judgments of similarity (or difference) are a function of shared and distinctive features. With nondirectional instructions, symmetry was found regardless of scale.

At present, Tversky's model of similarity judgments has yet to find strong support. However, the original findings of Tversky (1977) and Tversky and Gati (1978) on asymmetry still require explanation. Alternative ways to consider the cognitive processes involved in directional judgments of similarity are needed. On the other hand, our results suggest that a geometric model of cognition, based on nondirectional judgments which generally yield symmetrical distances, may be quite appropriate and may not involve any distortions of the way the judgments are made.

References

Baird, J. C., & Noma, E. (1978). Fundamentals of scaling and psychophysics. New York: John Wiley and Sons.

Bauer, C. L., & Fink, E. L. (1983). Fitting equations with power transformations: Examining variables with error. In R. N. Bostrom (Ed.), Communication yearbook 7 (pp. 146-199). Beverly Hills, CA: Sage.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. Journal of General Psychology, 49, 229-240.

Bower, G. H., Clark, M. C., Lesgold, A. M., & Winzenz, D. (1969). Hierarchical retrieval schemes in recall of categorized word lists. Journal of Verbal Learning and Verbal Behavior, 8, 323-343.

Edison, N. G. (1976). The effects of situational definition and time on the dyadic interactional process. Unpublished master's thesis, Michigan State University, East Lansing, Michigan.

Lakoff, G., & Johnson, M. (1980). Metaphors we live by. Chicago: University of Chicago Press.

Lorayne, H., & Lucas, J. (1974). The memory book. New York: Stein and Day.

Marron, T. (1985). Ambiguity and the predictability of the violation of the triangle inequality. Unpublished doctoral dissertation, University of Maryland at College Park, Maryland.

Ortony, A. (1979). Beyond literal similarity. Psychological

Review, 86, 161-180.

Simmon, S. (1955). The wise men of Helm and their merry tales. New York: Behrman House.

Smith, E. E., & Medin, D. L. (1981). Categories and concepts. Cambridge, MA: Harvard University Press.

Stein, B. S. (1977). The effects of cue-target uniqueness on cued recall performance. Memory and Cognition, 5, 319-322.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. Psychological Review, 69, 344-354.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. Journal of Verbal Learning and Verbal Behavior, 5, 381-391.

Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. Psychological Review, 80, 352-373.

Tversky, A. (1977). Features of similarity. Psychological Review, 84, 327-350.

Tversky, A., & Gati, I. (1978). Studies of similarity. In E. Rosch & B. B. Lloyd (Eds.), Cognition and categorization (pp. 79-98). Hillsdale, NJ: Lawrence Erlbaum Associates.

Woelfel, J., & Fink, E. L. (1980). The measurement of communication processes: Galileo theory and method. New York: Academic Press.

Footnotes

(1) Galileo, for example, a geometric model described in Woelfel and Fink, 1980, does not make all these assumptions.

(2) The effect of whether features were asked about before or after the similarity or distance judgments was generally insignificant. Limiting our analysis to the two pairs of countries whose features were inquired about in the questionnaire (U.S.S.R.-Syria and England-Ireland), we find only one result to be significant: Subjects who estimated difference using a closed-ended scale and for whom features were measured last saw U.S.S.R.-Syria to be significantly more different than those subjects who used a closed-ended scale and for whom features were measured first. In addition, for these two pairs, there were no significant interaction effects of order (whether the more prominent country was first or second in the comparison) by whether features were measured before or after the similarity judgments.

(3) Recall that the actual question put to subjects was not the prominence of the country, but which country they preferred to use as the referent when making comparison judgments.

Table 1

Proportion of subjects selecting country p as more prominent than prominent country q, by pairs of countries, based on pilot study

Pair #	p	q	Proportion(a)
1	U.S.A.	MEXICO	0.846
2	U.S.S.R.	POLAND	0.885
3	CHINA	ALBANIA	0.769
4	U.S.A.	ISRAEL	0.923
5	JAPAN	PHILIPPINES	0.808
6	U.S.A.	CANADA	0.923
7	U.S.S.R.	ISRAEL	0.615
8	ENGLAND	IRELAND	0.769
9	W. GERMANY	AUSTRIA	0.577
10	U.S.S.R.	FRANCE	0.577
11	BELGIUM	LUXEMBOURG	0.769
12	U.S.A.	U.S.S.R.	0.577
13	CHINA	N. KOREA	0.846

Table 1 continued

14	*INDIA*	CEYLON	0.769
15	*U.S.A.*	FRANCE	0.923
16	*U.S.S.R.*	CUBA	0.962
17	*ENGLAND*	JORDAN	0.615
18	*FRANCE*	ISRAEL	0.654
19	*U.S.A.*	W. GERMANY	0.846
20	*U.S.S.R.*	SYRIA	0.808
21	*FRANCE*	ALGERIA	0.731

(a) A proportion $\geq$ .654 is significant at $p < .05$; $N = 26$. The more prominent country in each pair is italicized. Subjects were asked to indicate which country they preferred to use as the referent of a comparison. The country so designated is considered the more prominent.

Table 2

Means (and standard deviations) of closed-ended scale of judgments of similarity or difference when more prominent country used as subject (SUBJ) versus more prominent country used as referent (REF)

	SIMILARITY CONDITION				DIFFERENCE CONDITION			
	SUBJ	(SD)	REF	(SD)	SUBJ	(SD)	REF	(SD)
MEXICO-U.S.A.	6.97	(4.85)	8.60	(3.42)	12.97	(4.68)	11.36	(4.37)
U.S.S.R.-POLAND	13.17	(5.27)	12.42	(4.84)	7.61	(4.88)	7.47	(4.35)
CHINA-ALBANIA	5.26	(4.54)	5.94	(4.49)	13.21	(4.71)	12.57	(4.87)
U.S.A.-ISRAEL	9.09	(5.97)	8.69	(5.97)	11.36	(5.48)	12.11	(4.90)
PHILIPPINES-JAPAN	9.94	(4.75) *	12.86	(4.10)	9.28	(5.14)	10.44	(4.93)
CANADA-U.S.A.	15.69	(4.89)	17.29	(2.75)	5.19	(4.91)	4.22	(3.40)
U.S.S.R.-ISRAEL	4.83	(4.79)	5.23	(5.00)	15.86	(5.11)	15.22	(5.09)
ENGLAND-IRELAND	13.51	(5.09)	13.19	(4.70)	7.81	(4.45)	8.39	(4.86)
AUSTRIA-W. GERMANY	13.31	(4.41)	14.60	(4.39)	7.17	(4.29)	6.83	(4.79)
FRANCE-U.S.S.R.	5.44	(3.38)	6.69	(4.85)	13.83	(5.05)	13.33	(5.11)
BELGIUM-LUXEMBOURG	14.46	(3.65)	13.03	(4.87)	5.94	(3.70)	5.66	(3.65)
U.S.A.-U.S.S.R.	5.66	(5.23)	5.47	(4.11)	16.14	(5.03)	15.94	(5.37)
CHINA-N. KOREA	13.71	(4.97)	13.00	(4.37)	8.50	(5.29)	6.86	(5.16)
CEYLON-INDIA	9.45	(5.11) *	12.29	(5.28)	8.63	(3.70)	7.67	(3.90)
U.S.A.-FRANCE	13.63	(3.08) *	10.78	(4.30)	7.83	(3.54)	8.89	(4.27)
CUBA-U.S.S.R.	13.97	(4.14)	15.74	(3.40)	7.08	(5.43)	7.19	(4.80)
JORDAN-ENGLAND	5.83	(3.59)	6.26	(4.73)	11.51	(4.31)	12.78	(4.32)

	SIMILARITY CONDITION				DIFFERENCE CONDITION			
ISRAEL-*FRANCE*	7.56	(4.48)	6.97	(4.48)	11.47	(3.68)	11.72	(4.30)
W. GERMANY-*U.S.A.*	10.22	(5.31)	11.86	(5.55)	10.39	(4.98)	10.39	(4.92)
SYRIA-*U.S.S.R.*	10.89	(5.41)	10.49	(4.97)	10.43	(4.87)	11.49	(4.33)
FRANCE-ALGERIA	8.11	(4.59)	8.34	(5.10)	12.62	(4.55)	11.11	(4.44)

Sample sizes range from 31 to 36 per pair. More prominent country is italicized.

* Asterisks between columns indicate a significant difference between the subject and the referent for a particular pair of countries ($p<.05$, two-tailed test).

(a) Means of transformed data such that X*= LOGARITHM (X+1), where X is the original reported distance and X* is the transformed score. Across subjects both countries (in any given pair of countries) appeared an approximately equal number of times as subject and as referent.

Table 3

Means (and standard deviations) for transformed judgments of similarity or difference when more prominent country used as subject (SUBJ) versus more prominent country used as referent (REF) in the magnitude scaling conditions

	SIMILARITY CONDITION				DIFFERENCE CONDITION			
	SUBJ	(SD)	REF	(SD)	SUBJ	(SD)	REF	(SD)
MEXICO-U.S.A.	4.78	(1.26)	4.18	(1.57)	4.84	(1.20)	4.69	(.79)
U.S.S.R.-POLAND	5.53	(1.89)	5.08	(1.45)	3.97	(1.61)	3.50	(1.55)
CHINA-ALBANIA	3.05	(2.11)	3.22	(2.02)	5.18	(.68) *	4.64	(1.24)
U.S.A.-ISRAEL	3.88	(1.88)	4.13	(1.40)	5.24	(.84)	4.84	(1.08)
PHILIPPINES-JAPAN	4.98	(1.30)	4.53	(1.30)	3.96	(1.55)	4.37	(1.18)
CANADA-U.S.A.	6.04	(.77)	6.00	(1.67)	2.94	(1.66)	3.15	(1.59)
U.S.S.R.-ISRAEL	3.04	(2.04)	3.38	(1.75)	5.25	(1.49)	5.18	(1.50)
ENGLAND-IRELAND	5.36	(1.69)	5.68	(1.09)	4.10	(1.39) *	3.28	(1.65)
AUSTRIA-W. GERMANY	5.68	(.97)	5.31	(1.78)	2.97	(1.82)	3.61	(1.88)
FRANCE-U.S.S.R.	3.81	(1.29)	3.52	(1.73)	4.96	(1.07)	5.21	(.98)
BELGIUM-LUXEMBOURG	4.79	(1.99)	5.30	(1.28)	2.87	(2.07)	2.36	(1.86)
U.S.A.-U.S.S.R.	3.90	(2.67)	3.23	(2.38)	5.94	(2.11)	5.89	(2.44)
CHINA-N. KOREA	5.20	(1.60)	5.66	(.80)	4.25	(1.14) *	3.74	(1.64)
CEYLON-INDIA	4.73	(1.86)	4.50	(1.39)	3.26	(1.66)	3.79	(1.70)
U.S.A.-FRANCE	4.92	(1.40)	5.21	(.83)	4.54	(.73)	4.20	(.92)
CUBA-U.S.S.R.	5.18	(1.84)	5.21	(1.89)	3.46	(1.95)	3.58	(1.68)
JORDAN-ENGLAND	3.54	(1.82)	3.88	(1.91)	4.76	(.82)	5.08	(.63)

	SIMILARITY CONDITION				DIFFERENCE CONDITION			
ISRAEL-*FRANCE*	3.93	(1.52)	4.06	(1.47)	4.64	(.92)	5.08	(.73)
W. GERMANY-*U.S.A.*	4.78	(1.14)	4.64	(1.80)	4.39	(.99)	4.56	(1.38)
SYRIA-*U.S.S.R.*	4.24	(1.51)	4.63	(1.75)	4.25	(1.64)	4.51	(1.31)
FRANCE-ALGERIA	4.06	(1.61)	4.07	(1.73)	4.76	(1.12)	4.81	(.95)

Sample sizes range from 32 to 37 per pair. More prominent country is italicized.

* Asterisks between columns indicate a significant difference between the subject and the referent for a particular pair of countries ($p<.05$, two-tailed test).

(a) Means of transformed data such that X*= LOGARITHM (X+1), where X is the original reported distance and X* is the transformed score. Across subjects both countries (in any given pair of countries) appeared an approximately equal number of times as subject and as referent.

Table 4

t-tests for asymmetry in the magnitude and closed-ended scaling conditions for subjects' estimates of similarity (SIM) and difference (DIF)

	Magnitude scaling		Closed-ended scaling	
	SIM	DIF	SIM	DIF
t =	.00	.24	1.48	-.77
p =	>.20	>.20	>.10	>.20

Note: df = 20 for all four (two-tailed) t-tests.

Table 5
Regression coefficients (and standard errors) for predictor variables on the pair U.S.S.R.-Syria, for similarity and difference(a)

	1-20 Scale				Transformed Magnitude Scale			
Predictor	Similarity	(S.E.)	Difference	(S.E.)	Similarity	(S.E.)	Difference	(S.E.)
(R ∩ S)	1.20	(1.37)	-.50	(1.49)	.29	(0.49)	-.75	(0.53)
(R - S)	2.86	(2.84)	-1.96	(2.82)	-.42	(0.89)	.75	(0.81)
(S - R)	-.36	(2.20)	-2.30	(2.34)	.20	(0.79)	.70	(0.82)
(order)*(R-S)	-.23	(2.71)	3.96	(2.65)	-.23	(0.96)	.08	(0.96)
(1 - order)*(S-R)	-1.21	(3.49)	5.12	(3.18)	-.08	(1.20)	-.25	(1.08)
Constant	7.10		11.87		4.70		3.94	
R-squared	.06		.06		.04		.13	
N	57		47		50		48	

(a) Order = 1 when the U.S.S.R. is the subject of the comparison; Order = 0 when Syria is used as the subject.
(R ∩ S) = the number of shared features of the U.S.S.R.-Syria.
(R - S) = the number of distinctive features of the U.S.S.R.
(S - R) = the number of distinctive features of Syria.

Table 6
Regression coefficients (and standard errors) for predictor variables on the pair England-Ireland, for similarity and difference(a)

	1-20 Scale				Transformed Magnitude Scale			
Predictor	Similarity	(S.E.)	Difference	(S.E.)	Similarity	(S.E.)	Difference	(S.E.)
(E ∩ I)	-1.86	(0.89)	-1.69	(2.09)	1.07*	(0.54)	-1.38*	(0.63)
(E - I)	-3.72	(2.79)	-.16	(1.90)	-.25	(0.67)	1.06	(1.10)
(I - E)	-.56	(2.04)	-1.61	(2.44)	-1.16	(1.08)	.86	(0.70)
(1 - order)*(E-I)	4.24	(3.25)	1.25	(2.84)	1.49	(1.09)	-1.02	(1.24)
(order)*(I-E)	3.87	(3.24)	2.70	(3.28)	1.12	(1.20)	-1.15	(1.28)
Constant	10.63		10.06		4.13		4.81	
R-squared	.05		.05		.15		.13	
N	57		47		50		48	

(a) Order = 0 when the England is used as the subject of the comparison; Order = 1 when the Ireland is used as the subject.
(E ∩ I) = the number of shared features of England-Ireland.
(E - I) = the number of distinctive features of England.
(I - E) = the number of distinctive features of Ireland.
* An asterisk indicates significance ($p < .05$) for R-squared or for a coefficient.