Integrating verbal quantitative information

HARRY M. HERSH and ALFONSO CARAMAZZA
Johns Hopkins University, Baltimore, Maryland 21218

An experiment is reported where we explore whether people can effectively integrate quantitatively vague information. Subjects were first asked to assign numerical values to the terms *much, slightly, larger,* and *smaller.* In a separate session 1 week later, they were asked to solve linear syllogism problems such as *A is slightly larger than B, C is much smaller than B, therefore A is . . . units distant from C.* Results indicate that the mean numerical values assigned to the terms in the syllogisms can be used to predict solutions to the syllogism problems, suggesting that subjects can integrate verbal quantitative information.

If B is much larger than A, whereas C is slightly larger than A, then surely B is larger than C. But how much larger? Do we use verbal hedges like *much* and *slightly* with sufficient precision to permit quantification of their verbal meanings?

There have been several attempts to quantify such semantic relationships (e.g., Cliff, 1959; Jones & Thurstone, 1955; Mosier, 1941; Osgood, 1952; Osgood, Suci, & Tannenbaum, 1957). Mosier hypothesized that the meaning of a word may be considered to contain two components, a constant component reflecting the overall meaning value along a continuum and a variable component representing the variation in the meaning of the word due to context, speaker, etc. The model assumes a unidimensional continuum along which every word must fall. The model also predicts, for example, that context effects will be independent of the word used and the individual involved. Although these and other simplifications reduce the model's explanatory and predictive power considerably, it was probably the first significant attempt to quantify meaning.

In an empirical investigation of his model, Mosier had subjects rate a list of evaluative adjectives (e.g., excellent, unsatisfactory) along an unfavorable-neutral-favorable continuum. He then scaled the responses by the method of successive intervals. The scale value for each word was interpreted as the constant component of the meaning of the word, while the spread of the distribution was considered as ambiguity, or the variable component. The data tended to support his model in that it was possible to assign each word to a scale value along the unidimensional continuum, and the variation in meaning about this scale value was approximately normally distributed. [These results were later supported by

Jones & Thurstone (1955) for a set of preference words and phrases.]

Cliff (1959) extended our understanding of the quantification of language by exploring the way in which adverbs operate on concepts. He proposed a formal model where adverbs functioned as multiplicative constants: That is, if an adjective can be described as a point on a scale, then the addition of an adverb would tend to change the scale position of the adjective in proportion to the distance between the original position and the scale origin.

Cliff's empirical findings clearly confirmed why combinations such as *unusually average* appear so awkward: The scale value of *average* is approximately zero, so multiplying the scale value of *average* by any multiplicative constant will leave the value unchanged—the adverb is superfluous.

Both Cliff and Mosier have shown that evaluative adjectives, alone and in combination with adverbs, can be represented as points on a continuum. In this paper we attempt a treatment of relational adjectives (e.g., larger, better) within this latter framework. The hypothesis suggested here is that, since evaluative adjectives can be described as points on a scale, relational adjectives should be described as distances between points on the scale. These distances, like the scale points, can in turn be modified through the use of adverbs. Extending the model further, if a relation can be considered in terms of a distance on a scale, then, when several relations are concatenated, as in a linear syllogism (A is taller than B, B is taller than C, etc.), the distances should be expected to combine similarly.

METHOD

Subjects

Sixty undergraduates at The Johns Hopkins University served as subjects in order to fulfill a requirement in an introductory psychology course.

Procedure

The study consisted of two experimental sessions. In the first session the subjects were given relational statements such as "A is slightly larger than B." The relational term consisted

of either *larger* or *smaller* and was modified by either *much* or *slightly*, or was left unmodified. One of the letters in each statement was designated as the midpoint of a scale from 1 to 100. The subjects' task was to determine where on the scale the other point should be placed. Subjects were presented with all six combinations of relational term and modifier. This session lasted approximately 2 min.

In the second session 1 week later, the same subjects were given a set of 36 modified linear syllogisms of the form

A is slightly larger than B
C is much smaller than B.

One of the three items in each problem was designated as existing at the midpoint of a 1 to 100 scale. The subjects' task was to simply decide how far apart A and C were in terms of this scale. The subjects were allowed to work at their own pace. The premises for the syllogisms were taken directly from the statements rated in the first session. All combinations of relational term and syllogism type were used, resulting in 72 different problems. In addition, anchoring was also varied independently. The problems were randomly assigned to subjects, with the restriction that each subject receive one-third of his problems from each anchoring position and each problem be equally often represented across subjects. This second session lasted approximately 25 min.

RESULTS AND DISCUSSION

Frequency distributions were obtained for each combination of relational term and hedge (adverb) used in the first session. The means and standard deviations of the absolute distances between the points are contained in Table 1. Since there were no significant rating differences between statements containing *larger* and statements containing *smaller*, the means of each statement pair were averaged to obtain a composite rating for each hedge. These means are also presented in Table 1.

A four-factor repeated measurement analysis of variance was performed on the distance ratings from the second session. There was an obvious effect for the variation in relational term, $F(8,72) = 619.99$, $p < .0001$. In addition, there were significant effects for anchoring and an interaction between the hedge used and the anchor, $F(2,18) = 78.65$, $p < .001$ and $F(16,144) = 24.85$, $p < .001$, respectively. The significant anchoring effects are due to some of the overall distances being greater than 50. When these distances were paired with an end anchor condition, the resulting scale value would be either greater than 100 or less than 0. The scale, however, ranged from 1 to 100, resulting in a clipping of the actual distances. (To avoid the possibility of any effects due to anchoring,

Table 1
Distances from the Center of the Rating Scale

Hedge	Smaller Mean	SD	Larger Mean	SD	Grand Mean
Slightly	5.29	4.54	5.78	4.37	5.53
(None)	17.81	7.50	19.34	8.49	18.57
Much	39.07	8.38	40.53	7.12	39.80

Table 2
Distances Between the Extreme Members of the Triads

Hedge Type*	Mean Distance Predicted	Actual
A is larger than B B is larger than C	37.14	29.80
A is slightly larger than B B is larger than C	24.10	23.70
A is larger than B B is slightly larger than C	24.10	20.40
A is slightly larger than B B is slightly larger than C	11.06	11.00
A is much larger than B B is larger than C	58.37	53.20
A is larger than B B is much larger than C	58.37	55.00
A is much larger than B B is much larger than C	79.60	76.80
A is slightly larger than B B is much larger than C	45.33	42.80
A is much larger than B B is slightly larger than C	45.33	42.90

Only one type of syllogism is shown, although there were eight variations.

only center anchored problems will be considered.)

Mean distances were obtained from each type of premise pair found in the syllogisms. Since there were no significant differences among the various types of syllogisms (i.e., A is larger than B or B is smaller than A), the means were then collapsed over syllogism type.

The syllogisms had been formed from statements used in the first session. Since the distance information in the two premises was integrated, it was assumed that the overall distance rating for each syllogism would reflect the concatenated ratings of the individual statements. Thus, the mean distance ratings for each statement in a syllogism were summed as an estimate of the overall rating. The results are shown in Table 2. As can be seen, the correspondence was quite close between the overall rating and the sum of the individual premise ratings.

While the analyses we have reported did not take into consideration the variability involved in the individual statements and in the overall ratings, the analysis of the means of the distributions was sufficient to demonstrate that subjects do integrate these vague relational statements in a quantitatively predictable manner. This result is especially interesting in light of current interest with the issue of how people use quantitatively vague information. Here the issue is, how is it that we can effectively communicate with such words as *few, slightly, very, sort of*, etc., when these words do not seem to specify a precise degree of quantification? The very fact that communication with these hedges is possible, however, should make it clear that there must be considerable interindividual agreement as to how to interpret these

vague terms. In addition, the fact that subjects can use this type of information in problem solving situations, such as the one we have reported, implies that we have adequate conceptual structures and mechanisms for handling the vagueness inherent in natural language. These two observations, that communication with vague terms is possible and that subjects can adequately use this type of information in problem solving situations, make a formal analysis of the quantitative meaning of such terms a much desired goal (see Hersh & Caramazza, 1975, for a preliminary attempt).

REFERENCES

CLIFF, N. Adverbs as multipliers. *Psychological Review*, 1959, **66**, 27-44.

HERSH, H. M., & CARAMAZZA, A. The quantification of vague terms. Paper presented at the Psychometric Society Meeting, Iowa City, April 1975.

JONES, L. V., & THURSTONE, L. L. The psychophysics of semantics: An experimental investigation. *Journal of Applied Psychology*, 1955, **39**, 31-36.

MOSIER, C. I. A psychometric study of meaning. *Journal of Social Psychology*, 1941, **13**, 123-140.

OSGOOD, C. E. The nature and measurement of meaning. *Psychological Bulletin*, 1952, **49**, 197-237.

OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning*. Urbana: University of Illinois Press, 1957.