Journal of Experimental Psychology:
Human Perception and Performance
1983, Vol. 9, No. 3, 487–490

Observations

Numerical Reference Points Reexamined: A Reply to Shoben, Čech, and Schwanenflugel

Keith J. Holyoak
Human Performance Center, University of Michigan

Shoben, Čech, and Schwanenflugel (1983) criticized the distance ratio model of numerical comparisons proposed by Holyoak (1978) and presented an alternative subtraction model. I argue that the empirical evidence obtained by Shoben et al. is inadequate to reject the distance ratio model. Furthermore, their subtraction model is theoretically and empirically weak as an account of numerical comparisons and cannot be readily generalized to other types of symbolic comparative judgments. Discriminability models, of which the distance ratio model is a special case, offer a more promising theoretical framework for investigating the influence of reference points in a variety of judgment tasks that can be performed with a broad range of symbolic continua.

Symbolic comparative judgments are influenced by stimuli that function as reference points. In particular, differences between stimuli with magnitudes close to that of a reference point appear to be more discriminable than are differences between stimuli with magnitudes far from it. The "semantic congruity effect" typically observed in comparative judgment tasks may in part be a special case of this general principle. Some time ago I investigated the possibility that explicit reference points will influence judgment time in a numerical comparison task (Holyoak, 1978). Subjects were timed as they decided which of two digits was closer to (or further from) a third digit (the reference point) in numerical magnitude. As predicted, an effect of distance to the reference digit was observed. Holding the pair constant, reaction time (RT) increased with distance to the reference digit. For example, subjects could select the closer member of the pair 3, 4 more quickly when the reference digit was 6 rather than 7. I interpreted this result as evidence that discriminability is increased in the vicinity of a reference point.

In addition to this effect of distance to the reference point, I obtained an effect of the difference between the two stimulus-to-reference distances: the larger the difference, the faster the decision. For example, for the reference digit 5, subjects were faster to choose the closer member of pair 2, 4 than of pair 3, 4. The latter result is essentially the same as the "distance effect" obtained in many studies of comparative judgments with stimulus pairs (Moyer & Landauer, 1967). Following a suggestion made by Jamieson and Petrusic (1975), I assessed the possibility that subjects were judging the ratio of differences. In several experiments, the RT pattern was described accurately by Equation 1,

$$RT_{ij(RP)} = \alpha(s_i - s_{RP})/(s_j - s_{RP}) + K, \quad (1)$$

where $RT_{ij(RP)}$ is the reaction time to choose which of stimuli i and j is closer to reference point RP, and s_i is the magnitude value of the stimulus closer to the reference point RP. This "distance ratio" predicts the effects of both the distance from the comparison stimuli to the reference point and the difference between the two distances.

An intriguing outcome of these tests of the distance ratio was that the optimal scale for measuring distance differed depending on whether the triplet was *unilateral* (both digits either larger or smaller than the reference point, such as 7, 9 for reference point 5) or *bilateral* (digits straddling the reference point, such as 3, 9 for reference point 5). For unilateral triplets the optimal scale was logarithmic (or some similar negatively accelerated function), whereas for bilateral triplets it was linear. Bilateral triplets also tended to produce slower RTs overall. These scale differences were very robust. In Experiment 2, for example, for unilateral triplets the ratio computed on a log scale accounted for 42% more variance than did the ratio computed on a linear scale, whereas for bi-

This article was completed while the author was supported by a National Institute of Mental Health Research Scientist Development Award (1-K02-MH00342-02).

Requests for reprints should be sent to Keith J. Holyoak, University of Michigan, Human Performance Center, 330 Packard Road, Ann Arbor, Michigan 48104.

lateral triplets the linear form of the ratio accounted for 58% more variance than did the logarithmic form. I interpreted these results as evidence that subjects based decisions about unilateral triplets on subjective digit magnitudes, which a great deal of evidence indicates are negatively accelerated (e.g., 2 and 3 seem farther apart than do 8 and 9). In contrast, decisions about bilateral triplets require a subtraction process based on a linear scale because subjective magnitudes may yield the wrong answer (e.g, on a log scale 8 is actually closer to 5 than 3 is).

Rebuttal to Shoben et al.'s Critique

Shoben, Čech, and Schwanenflugel (1983) have criticized the conclusions drawn in my 1978 article (although they replicated its major findings). They argue that an assumption of the distance ratio model is incorrect, and they propose a new model, which they contend gives a "somewhat better" account of the relevant data. I address these two claims in turn.

Identification of Triplet Type

In my 1978 article, I assumed that the first stage in the decision process is to assess whether the given triplet is unilateral or bilateral:

> The first stage is necessary if subjects actually use different distance scales for unilateral vs. bilateral triplets. Presumably it should be easier to identify on which side of the reference point a digit falls if it is relatively far from the reference point. However, such an effect will tend to be obscured by the opposite effect of distance from reference point predicted for later stages. . . . (p. 212)

In terms of Equation 1, I assumed that the estimated value of α would be reduced by some amount due to the inverse relationship between distance to reference point and time to identify triplet type.

In their Experiment 1, Shoben et al. (1983) confirmed my assumption that time to identify triplet type decreases with the distance between the pair and the reference point by having subjects overtly decide whether triplets were unilateral or bilateral. Shoben et al. then argued that the distance ratio model must predict that RT will increase more with distance to the reference point when subjects know the triplet type in advance than when they do not because such foreknowledge would obviate the need to make a decision about triplet type, and hence remove the stage in which RT decreases with distance to reference point. (They refer to this as a "revised" distance ratio model, but in fact it is simply the original.) In a comparison across their Experiments 2 and 3, Shoben et al. found no support for this putative prediction. Informing the subjects of the triplet type (unilateral or bilateral) 2 sec in advance of the stimuli did not increase the effect of distance to the reference point, although the cue did tend to reduce overall RT.

Shoben et al.'s study highlights the need to clarify the process by which triplet type might be judged in the context of the distance ratio model. My 1978 formulation of the model treated this stage in an incidental fashion, concentrating theoretical focus on subsequent processing stages (generating and comparing the stimuli-to-reference distances). However, one can question Shoben et al.'s identification of the process with the explicit unilateral/bilateral judgment task used in their Experiment 1. Shoben et al. (1983) told their subjects that "they would have to decide whether two numbers were both numerically larger or smaller than a third number described as the reference point, or whether one number was larger and one was smaller than the reference point" (p. 230), and to press an appropriate response key. In contrast, when subjects are simply asked to choose which of two numbers is numerically closer to a third, there is no explicit mention of triplet type at all, far less of explicit definitions in terms of "larger" and "smaller." There is also, of course, no requirement for an overt, conscious response based on triplet type. In the latter task, a plausible possibility is that subjects immediately begin to assess subjective differences and compare them, using an iterative procedure such as a random walk. If the two digit-to-reference differences are consistently of opposite sign, this may trigger a strategy shift to use of subtraction. Bilateral triplets would therefore be evaluated in a different manner than would unilateral triplets without triplet type necessarily being consciously identified as such. Furthermore, evaluation of triplet type would be made partially in parallel with generation and comparison of digit-to-reference distances.

There is no reason to assume, then, that the process of consciously evaluating triplet type, as studied by Shoben et al., must yield an RT pattern comparable to the more tacit process that may underlie assessment of distance ratios. Similarly, Shoben et al.'s failure to detect any change in the effect of distance to the reference point as a function of precuing with triplet type is not definitive. An obvious possibility is that their subjects were simply unable or unwilling to make effective use of this information in the 2 sec prior to presentation of the triplet. At least for the relatively unpracticed subjects used by Shoben et al., it may have been easier to simply perform the usual implicit decision process rather than to consciously adopt a particular strategy as a function of the cue.

A stronger manipulation than Shoben et al.'s precuing procedure would be to examine performance for the same set of unilateral triplets with and without intermixed bilateral triplets. If subjects receive only one type of triplet, they should eventually cease to evaluate triplet type, even tacitly. In that case RT may, as Shoben et al. suggested, no longer increase as much with distance to the reference point. Shoben et al. failed to note that I performed this test (Holyoak, 1978, Table 5, p. 227) and obtained the predicted result. With intermixed bilateral triplets, RT for the unilateral triplets increased with distance to the reference digit. Without bilateral triplets, RT for the same unilateral cases actually decreased with distance. Although I interpreted this result in a somewhat different fashion, it can certainly be viewed as evidence that triplet type influences judgment strategies.

Shoben et al.'s Subtraction Model

Even though the empirical evidence Shoben et al. directed against the distance ratio model is inconclusive, their criticism would be strengthened if their own model offered an attractive alternative. However, this is not the case. Their model assumes that subjects choose the closer of two digits by subtracting each from the reference point and then choosing the larger of the two differences. There is no need to even implicitly assess triplet type. Although at this general level of description the subtraction model seems parsimonious, closer examination alters this impression. The postulated subtraction process involves a number of ill-supported assumptions (e.g., subtraction time is hypothesized to be linearly related to the minimum of the difference and the lesser number, and to be increased by a constant time if a digit is larger than the reference point). The latter assumption is apparently required to account for the fact that for unilateral pairs only, RT is longer when the pair is greater than the reference point (Holyoak, 1978). This pattern is necessarily predicted by the distance ratio model if subjects are in fact assessing the distance ratio on a logarithmic measurement scale.

Shoben et al. derived quantitative fits for their model, using the data of their Experiments 2 and 3 and Experiment 2 of Holyoak (1978). In terms of overall variance accounted for, the subtraction model and the distance ratio model appear to differ very little. However, when fitted separately to data for unilateral and bilateral triplets, the subtraction model allows three free parameters whereas the distance ratio model allows only two. Shoben et al.'s treatment of the data from my experiment is especially unconvincing. They need to add an entirely different process to account for the data obtained for two particular reference points, and they fail to note that the distance ratio model produced a multiple correlation of .91 for this entire data set (both unilateral and bilateral cases) using just three free parameters (Holyoak, 1978), which is a degree of fit that is apparently at least equal to that of the subtraction model plus its ad hoc addition.

In any case, as Birnbaum (1973) has argued, correlations as measures of goodness of fit provide weak evidence for choosing among alternative models. More analytic procedures involve searching for qualitative tests that differentiate the candidate models. Consider, for example, the predictions of the two models for two triplets with 5 as reference point: 4, 7, which is bilateral, and 6, 7, which is unilateral. For such cases, in which the magnitude difference is one, RT is substantially longer for the bilateral than for the unilateral triplet (Holyoak, 1978). The distance ratio model, which assumes that the measurement scale will differ as a function of triplet type, can account for the observed RT differences. But what does the subtraction model, as outlined by Shoben et al., predict for this example? For the bilateral case, the digit 4 yields an index of subtraction difficulty of 1 (the difference); the digit 7 yields an index of 2 (the difference) plus 1 because the digit is larger than the reference point. The overall subtraction index, which should predict RT, is the sum of the indices for the two digits (i.e., 4). For the unilateral case, the digit 6 yields an index of 1 (the difference) plus 1 because the digit exceeds the reference point; the digit 7 yields an index of 2 (the difference) plus 1; therefore, the total subtraction index is 5. Because the magnitude difference is equal across the two cases, the subtraction model erroneously predicts that RT will be longer for the unilateral than for the bilateral triplet—a qualitative violation of the model.

Overall Assessment

If the subtraction model has little to recommend it as an account of numerical comparisons, it is even less promising as a general model of magnitude comparisons. The basic phenomenon that motivated the development of reference point models—relatively fast RTs to make comparisons in the vicinity of a reference point—has been found for a variety of nonnumerical domains. These include judgments of similarity among semantic categories (Hutchinson & Lockhead, 1977) and among color concepts (te Linde & Paivio, 1979), as well as judgments of distance among geographical locations (Baum & Jonides, 1979). It is difficult to envisage how an explicit subtraction process might be used in such tasks. In addition, Holyoak and Mah (1982) have demonstrated that reference points influence not only decision time but also unspeeded geographical-distance judgments. They found that the rated distances between cities in the vicinity of a specified reference point (e.g., the Pacific Ocean) were expanded relative to the distances between cities far from it. The subtraction model falls in the category that Holyoak and Mah (1982) term "nondiscriminability" models, which predict that reference points can effect only the time to make speeded judgments and not the judgments themselves (in the absence of speed pressure). Accordingly, Holyoak and Mah's (1982) evidence against the adequacy of nondiscriminability models will apply to any extension of the subtraction model to unspeeded judgment tasks.

Although the subtraction model is not a promising alternative, I do not mean to imply that the distance ratio model as formulated by Holyoak (1978) provides an entirely adequate account of symbolic comparative judgments. As Holyoak and Mah (1982) have pointed out, not all observed performance patterns have been in accord with the precise quantitative form that the distance ratio predicts. We suggested that Equation 1 should be replaced by the more general formulation of Equation 2:

$$R_{ij(RP)} = J_{RP}(f[s_i - s_{RP}] - f[s_j - s_{RP}]), \quad (2)$$

where $R_{ij(RP)}$ is a response related to the subjective difference between stimuli i and j with respect to reference point RP; s_i, s_j, and s_{RP} are the subjective scale values of the two stimuli and the reference point, respectively; J_{RP} is a monotonic judgment function associated with reference point RP; and f is an increasing, negatively accelerated function. The effect of f is to expand differences between stimuli close to the reference point relative to differences between stimuli far from it. Equation 2 can be taken as an expression of the general class of discriminability models; the distance ratio model (Equation 1) can be viewed as a special case in which the response measure is RT, f is logarithmic, and J_{RP} is exponential. Within this more general theoretical framework, the influence of reference points on various comparative-judgment tasks involving different types of subjective continua remains a worthwhile direction for research.

References

Baum, D. R., & Jonides, J. Cognitive maps: An analysis of comparative judgments of distance. *Memory & Cognition,* 1979, *7,* 462–468.

Birnbaum, M. H. The devil rides again: Correlation as an index of fit. *Psychological Bulletin,* 1973, *79,* 239–242.

Holyoak, K. J. Comparative judgments with numerical reference points. *Cognitive Psychology,* 1978, *10,* 203–243.

Holyoak, K. J., & Mah, W. A. Cognitive reference points in judgments of symbolic magnitude. *Cognitive Psychology,* 1982, *14,* 328–352.

Hutchinson, J. W., & Lockhead, G. R. Similarity as distance: A structural principle for semantic memory. *Journal of Experimental Psychology: Human Learning and Memory,* 1977, *3,* 660–678.

Jamieson, D. G., & Petrusic, W. M. Relational judgments with remembered stimuli. *Perception & Psychophysics,* 1975, *18,* 373–378.

Moyer, R. S., & Landauer, T. K. Time required for judgments of numerical inequality. *Nature,* 1967, *215,* 1519–1520.

Shoben, E. J., Čech, C. G., & Schwanenflugel, P. J. The role of subtractions and comparisons in comparative judgments involving numerical reference points. *Journal of Experimental Psychology: Human Perception and Performance,* 1983, *9,* 226–241.

te Linde, J., & Paivio, A. Symbolic comparison of color similarity. *Memory & Cognition,* 1979, *7,* 141–148.

Received November 22, 1982 ■