A context-sensitive representation
of word meanings

HENRY M. HALFF, ANDREW ORTONY, and RICHARD C. ANDERSON
University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

This study provides a numerical representation of contextual effects on the meanings of words, constructed from the order judgments of 19 subjects concerning the word "red" in 19 sentences. Subjects judged whether or not the red object mentioned in a sentence was redder than, less red than, or could be equally as red as the red object mentioned in each of the other sentences. These judgments were well described as an interval order. This means that the red ascribed in a sentence can be represented by a real interval with judgments of equally red corresponding to overlapping intervals. Semiorder axioms were not met, indicating that the width of the interval varied from sentence to sentence. Possible ways of incorporating the result into theories of semantic memory were discussed, as well as ways of accounting for the pronounced individual differences which were observed.

This paper presents a formal representation of the effects of sentence context on word meaning. Our working hypothesis is that words do not have a few discrete, qualitatively distinct meanings. Instead they act in different contexts to establish boundaries in an underlying continuum of knowledge. Thus, "The fire engine is red" is not represented in memory solely as a proposition attributing the property of redness to a fire engine. Rather, the sentence also delimits the set of appropriate representations of the color of the fire engine. Redness attributed in another context might give rise to a different, nonoverlapping set of possibilities.

Persuasive a priori arguments that the meanings of terms in ordinary use are context sensitive have been set forth by Wittgenstein (1963). He claimed that the senses of words in different occurrences and contexts are best considered as related only by "family resemblance," since it is impossible to determine any fixed defining characteristics. A related conclusion has been reached by linguists such as Weinreich (1971) and Lakoff (1972) when addressing the semantic atomism implicit in interpretive semantics (e.g., Katz, 1966, 1972).

There are several lines of empirical evidence showing semantic flexibility (Anderson & Ortony, 1975; Barclay, Bransford, Franks, McCarrell, & Nitsch, 1974). Especially interesting is the work of Labov (1973), who has been concerned to discredit the "categorical view" of word meanings. This view, he says, implicitly defines linguistic units in terms of discrete, invariant categories, which are qualitatively distinct, conjunctively defined, and composed of atomic primitives. Labov's experi-

ments show the fuzziness of the concept cup. The width, height, contents, material of construction, and number of handles were all factors in determining whether a drawing of an object would be called a "cup," "mug," "bowl," "vase," etc. Labov found that the influence of any dimension on the tendency to call an object a cup systematically interacted with context, that is, the values of the object on other dimensions. He thus concluded that the rather vague and hedged definitions found in ordinary dictionaries—replete with "usually," "often," "sometimes"—are in much closer agreement with language use than the all-or-none definitions assumed in most linguistic and psychological models.

The research described here is an attempt to provide a numerical representation of the boundaries and ranges determined by particular words in various contexts. The choice of a numerical representation is motivated by our belief that many important cognitive relations possess all the properties of the real numbers. One should, for example, be able to fill the gap between any two possible cognitive representations of attributed redness in a fashion which is impossible with a countable set of entities. This view implies that linguistic representations, being at most countable, cannot be isomorphic with the knowledge or experience that they convey. Rather, language can only serve to map more or less unique parts of the underlying cognitive space. This view illuminates the relation between context and semantic flexibility. Context operates to narrow the possible representations of an utterance to an appropriate interval. In this way, the uniqueness or precision of linguistic representations is greatly extended. It is obvious, for example, that we can communicate a great deal about the color of an object in a particular context using only a very small set of color names. If semantic flexibility is thus related to the underlying continuity of knowledge, a numerical representation seems to be the most natural way of capturing the relationship.

Table 1
Sentences Used to Define Contexts

Anger	He saw red when his secretary came in an hour late.
Apple	Red apples were used in the salad.
Army	The Red Army is very well disciplined.
Beet	Red beets make a good garnish.
Blood	The knife blade was red from the blood of the victim.
Brick	The red brick house was being demolished.
Dress	Her red dress made her the center of attention.
Eyes	The laborer's eyes were red from the smoke and dust.
Face	His face turned red with embarrassment.
Fire	The fire is red hot.
Fire engine	The red fire engine raced down the street.
Flag	The U.S. flag is red, white, and blue.
Hair	The boy with red hair stood out in the crowd.
Ink	The teacher wrote her comments in red ink.
Light	He stopped his car when the stop light turned red.
Paint	The red paint spilled onto the floor.
Soil	The red soil of Alabama is good for raising cotton.
Sunburn	His skin was red due to sunburn.
Sunset	As the sun set the sky turned red.
Wine	They drank fine red wine with their meal.

The research concerned the effects of context on the representations of the word "red." The assumptions are that the use of a word such as "red" in some particular context conveys information defining a set of plausible cognitive representations, and that in many cases these representations can be fully ordered by the subject (perhaps in terms of their projection on a redness vector or distance from a prototype). Given these assumptions, it becomes possible to obtain a rich characterization of the different representations of a term by looking at simple order judgments. Subjects were asked, for example, if the red of a sunburn was definitely redder than, definitely less red than, or possibly equally red as the red of a fire engine. This procedure was repeated for all possible pairs drawn from a set of sentences, each of which used the word in a different context. It was assumed that a subject would choose the first alternative if each of his cognitive representations of sunburn red was redder than any of his representations of fire-engine red. He would choose the second alternative if all his fire-engine reds were redder than any sunburn red. If some fire-engine reds were as red as some sunburn reds, he would indicate this by selecting the third alternative.

These assumptions may be expressed formally as follows: If a, b, c, d, etc. represent different contexts, then we represent the three possible responses as aRb, bRa, and a$\sim$b for some pair of compared contexts, according to whether the subject decides that the red in a is redder than that in b, that the red in b is redder than that in a, or that they could be equally red. Thus, "R" and "$\sim$" represent dominance and indifference relations, respectively. Fishburn (1970, 1973) has dealt extensively with empirical systems characterized by dominance and intransitive indifference. He has shown that under many conditions such systems can be represented by real intervals if the following axiom holds:

$$aRb \ \& \ cRd \rightarrow aRd \text{ or } cRb. \qquad (1)$$

Systems in which Axiom 1 holds are called interval orders. Fishburn's representation theorem states that if Axiom 1 is true, there exist two real-valued functions, L and U > L, such that aRb if and only if L(a) > U(b). In the present case U and L represent, respectively, upper and lower bounds of redness in each context.

Assuming that Axiom 1 holds, one can examine the scaling solution for further information. If, for example, aRb and bRc implies dRc or aRd, the data can then be scaled as a semiorder (cf. Suppes & Zinnes, 1963) in which U − L is a constant. In this case, one could propose that the effect of context is to change the amount of redness conveyed by the term while leaving the range of redness, U − L, constant for all contexts. An even stronger condition could hold, one in which $\sim$ is an equivalence relation, that is, transitive, symmetric, and reflexive; if so, a scaling solution would exist with L = U. Were this the case, the concept of a range of redness would not be needed after all; each use of "red" could be represented as a point.

In summary, our hypothesis is that "red" results in different representations in different contexts, and that these variations can be captured in an interval order scale. Also investigated was the possibility of using the stronger axioms of a semiorder or full order.

METHOD

Subjects

The subjects were 19 students in an educational psychology course who participated in the experiment as part of a course requirement.

Materials

The materials were the 190 pairs that can be constructed from the 20 sentences in Table 1. Each of the sentences contained the word "red." The word itself was unqualified; phrases such as "very red" or "as red as an apple" were not included. The sentences were chosen from the productions of other students in such a way as to maximize the range of redness across sentences, as judged by the subjects who produced the sentences, and the popularity of the theme, as indicated by frequency of use by the subjects producing the sentences. "Red" was used in its literal sense in all but two of the sentences, the ones about anger and the Red Army.

Procedure

The original design called for 190 pairs formed from 20 sentences. Due to an error, however, one pair was omitted and another duplicated. This error was not discovered until the subjects had been tested. The solution to the problems this raised is discussed in the Results section. Each sentence was the first member in half the pairs in which it appeared and the second member in the other half. The 190 sentence pairs were randomly ordered and divided into 19 blocks. Each block was typed on a separate page; the pages were then collated into booklets in 19 different orders, forming a Latin square. Finally, a 20th page was added to each booklet. This last page contained 10 pairs randomly chosen from different blocks and was included as a check on the subject's reliability.

Each of the subjects was given a machine-readable answer sheet, a mimeographed set of instructions, and one of the booklets described above. The experimenter read the instruc-

Table 2
Consistency Statistics for Each Subject

Subject	r	Difference Error	Overlap Error	Average Error	Axiom Failure	Tau	One-Wide	Two-Wide
1	.724	.308	.143	.193	.006	.842	4	0
2	.894	.080	.130	.123	.001	.903	3	2
3	1.000	1.714	.463	.515	.014	.979	1	0
4	.756	.310	.549	.468	.025	.818	6	1
5	.876	.400	.227	.257	.004	.897	4	0
6	1.000	.520	.548	.544	.015	.873	7	1
7	.802	.877	.368	.538	.020	.814	6	2
8	.933	.805	.762	.772	.034	.839	5	1
9	.909	2.278	.784	.942	.039	.931	3	0
10	.028	3.600	4.184	4.082	.248	.874	3	1
11	.693	.296	.142	.181	.005	.935	1	0
12	.896	2.864	.470	.778	.024	.976	0	0
13	−.051	1.592	3.861	3.211	.237	.832	7	1
14	1.000	2.012	.626	.825	.028	.943	2	0
15	.530	1.083	.208	.269	.003	.949	2	0
16	1.000	.421	.518	.485	.028	.893	3	0
17	.724	.229	.911	.719	.048	.880	5	1
18	.745	.409	1.090	.912	.051	.769	8	0
19	.753	1.062	1.640	1.532	.091	.884	4	0
Mean	.748	.901*	.916*	.913	.047*	.886	3.90	.53

*Weighted by number of cases.

tions to the subjects as they followed along on their copies. The most relevant part of the instructions was as follows: "If you think that the red object in Sentence A is definitely redder than the red object in Sentence B, mark Alternative A on your answer sheet. If you think that the object in Sentence B is definitely the redder, mark Alternative B on the answer sheet. If you think that the two objects could be equally red, mark Alternative E. In making these judgments, remember that we are interested in what the description normally conveys to you. Hence, you should only consider the normal range of redness implied by the sentence." The subjects worked through the booklets at their own pace.

RESULTS

All subjects were first evaluated on reliability using the pairs duplicated on the last page of each booklet. Each of the duplicated pairs, (a, b), was assigned a score of −1 if aRb, 0 if a⌣b, and 1 if bRa. Then, for each subject, the Pearson product-moment correlation coefficient was computed between scores on the first and second occurrences of the duplicated items. These coefficients are presented in Table 2. The bulk of the subjects were not perfectly reliable. In fact, two of them appeared to be responding at random. However, we believed that reliability was good enough to allow a test of the interval-order axiom.

To deal with the procedural error mentioned above, the Dress context was arbitrarily eliminated from all further analyses, since the sentence pair Dress and Eyes was not included in the materials. The other alternative, eliminating the Eyes context was also examined, and the results were essentially the same in either case.

The procedure suggested by Fishburn's (1970) proof of the representation theorem is ranking the set of values, V, consisting of all U(a)s and all L(a)s. To see how this can be done, consider an example involving the

four contexts, *eyes*, *brick*, *ink*, and *apple*. Suppose that a subject indicated that eyes could be as red as brick but is redder than both ink and apple, that brick could be as red as ink but is redder than apple, and that ink could be as red as apple. Denoting each context by its initial letter, we have

$$e⌣b, eRi, eRa, b⌣i, bRa, i⌣a.$$

To rank the eight points in V, we need only count the number of points exceeded by each member of V. Consider first the number of points below L(b), the lower limit of brick. Since brick is redder than apple, L(b) must exceed both the upper and lower bounds of apple. Also, since brick is redder than apple, and apple could be as red as ink, L(b) > U(a) ⩾ L(i), so that L(b) therefore exceeds three points in V, namely, U(a), L(a), and L(i). Now consider the points falling below U(b), the upper bound of brick. Naturally, those points falling below L(b)—U(a), L(a), and L(i)—will also fall below U(b). We may also assume that U(b) exceeds[1] the lower bounds of those objects which could be red as brick; that is, U(b) also exceeds its own lower bound, L(b), and that of ink, L(i). Finally, note that brick would be as red as eyes, and eyes are redder than ink. This implies that U(b) ⩾ L(e) > U(i); hence, U(b) also exceeds the upper bound of ink. U(b) therefore exceeds six points in V, namely, U(a), L(a), L(i), L(b), L(e), and U(i). This line of reasoning may be developed into the following general algorithm for determining the number of points exceeded by any point in V. We first introduce some notation:

$N_L(a)$ is the number of contexts, b, such that aRb.
$N_U(a)$ is the number of contexts, b, such that not (bRa).

Now consider the set of points that are ranked below L(a). Clearly, L(a) > U(b) if and only if aRb. We will assume that L(a) > L(b) if aRb, or if $a \frown b$ and $N_L(a)$ > $N_L(b)$. This latter condition [$N_L(a) > N_L(b)$] indicates there is a context c such that aRc and not (bRc); such a context could exist only if L(a) > L(b). Next, consider the set of points ranked below U(a). Again, U(a) > L(b) if and only if not (bRa). We will further assume that U(a) > U(b) if aRb, or if $a \frown b$ and $N_U(a)$ > $N_U(b)$. This latter condition [$N_U(a) > N_U(b)$] indicates there is a context c such that cRb and not (cRa); such a context could exist only if U(a) > U(b). Now let $M_L(a)$ be the number of points such that $a \frown b$ and $N_L(a)$ > $N_L(b)$, and let $M_U(a)$ be the number of points such that $a \frown b$ and $N_U(a)$ > $N_U(b)$. Then the number of points in V exceeded by L(a) is $2N_L(a) + M_L(a)$, and the number of points in V exceeded by U(a) is $N_U(a) + N_L(a) + M_U(a)$.

Applying this algorithm to the above example, we see that $N_L(b)$ = 1 (because bRa), and $M_L(b)$ = 1 [because $b \frown i$, and $N_L(b) = N_L(i) + 1$], $N_U(b)$ = 4 (because $b \frown c$, $b \frown b$, $b \frown i$, and bRa), and $M_U(b)$ = 1 [because $b \frown i$, and $N_U(b) = N_U(i) + 1$]. Thus, $2N_L(b) + M_L(b)$ = 3, and $N_U(b) + N_L(b) + M_U(b)$ = 6. Applying the algorithm to the other points in V reveals that the numbers of points exceeded by L(e), L(b), L(i), and L(a) are 5, 3, 0, and 0, respectively; corresponding figures for U(e), U(b), U(i), and U(a) are 6, 6, 4, and 2, respectively. Using these figures to rank the points in V produces the following result:

$$U(e) = U(b) > L(e) > U(i) > L(b) > U(a) > L(i) = L(a).$$

For each subject, the points in V were ranked according to the scheme just outlined, and each point was assigned its rank order, with ties being assigned the average of tied ranks. A graphical presentation of the results of this analysis for Subject 1 appears in Figure 1.

The implication of Fishburn's proof is that if Equation 1 is satisfied, the scaling procedure that has just been described will yield U and L such that $a \frown b$ if and only if the intervals [L(a), U(a)] and [L(b), U(b)] overlap. Failure of the axiom could result in two types of inconsistency. First, pairs with overlapping intervals could be judged different, and, second, pairs with nonoverlapping intervals could be judged indifferent. To evaluate the extent of the first type of error, called "overlap error," each pair in which aRb was assigned an overlap measure of 0 if L(a) > U(b), or U(a) − L(b) + 1 if L(a) < U(b). To evaluate the second type of inconsistency, "difference error," each pair in which $a \frown b$ was assigned a distance measure of max [0, L(a) − U(b), L(b) − U(a)]. The average overlap of pairs judged different, the average distance between pairs judged indifferent, and the weighted average of these two statistics was then computed. Finally, the overall proportion of occasions on which Equation 1 failed, called

"axiom failure," was computed as a proportion of the number of times the premise was met and a ≠ c and b ≠ d. The four types of error statistics are listed in Table 2. Although there is no absolute standard to judge these statistics by, they all seem quite low. Notice that scaling consistencies bear an obvious relation to the reliability of the subjects' judgments. When Subjects 10 and 13 are discounted, average error is about three-fifths of an ordinal unit and the scaling axiom fails in only .025 of the possible cases.

Having shown that Axiom 1 fits reasonably well, we will not consider stronger representations. A semi-order can be scaled in such a way that U − L is a constant. Although the scaling solution used here does not guarantee a semiorder representation if one exists, it does uniquely order the points in V. Hence, a sufficient condition for a semiorder is that U(a) > U(b) if and only if L(a) > L(b). To examine this, we computed Kendall's tau between U and L for each subject. The results are shown in Table 2. Some of the taus are quite high, suggesting the possibility that these subjects' data may be described by a semiorder. These statistics, however, are somewhat deceptive in that they entail the number of *pairs* that are involved in violations of the semiorder condition as opposed to the number of *individual contexts* that are involved in such violations.

Another way of viewing the data is to determine the percentage of contexts whose redness interval is wider than the interval of some other context. We will say that a context, a, is one-wide if there is another context, b, such that L(a) < L(b) and U(a) > U(b); a is a two-wide context if there are contexts b and c such that L(a) < L(b), U(b) < L(c), and U(a) > U(c). Thus, the one-wide contexts for Subject 1 are fire, sunburn, hair, and blood. (Subject 1 has no two-wide contexts.) The number of one-wide and two-wide contexts for each subject are shown in Table 1. If Subjects 10 and 13 are excluded, the overall percentages of one- and two-wide contexts are 19.8 and 2.5, respectively. Neither one-wide nor two-wide contexts are representable in a semi-
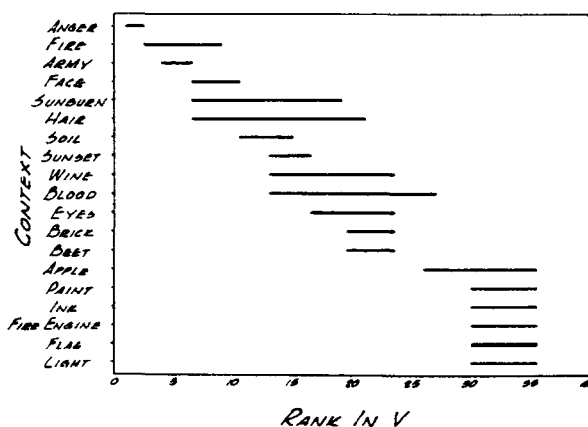

Figure 1. The intervals [U(a), L(a)] for each context, a, as determined by the data from Subject 1.

order. It would appear, therefore, that a semiorder characterization can be rejected; however, a definite conclusion is not possible because of the lack of reliability of the judgments.

Finally, it is interesting to note the remarkable lack of agreement among subjects. Kendall coefficients of concordance between U, L, and U − L were .08, .14, and .20, respectively. Another statistic which gives some idea of the degree of concordance is the extent to which one subject decides aRb when another decides bRa. For each pair of subjects, we computed the number of times that aRb for one subject *and* bRa for the other, and the number of times that aRb for one subject *or* bRa for the other. The overall ratio of the former number to the latter was .10. Neither this statistic nor the coefficients of concordance were substantially affected by eliminating Subjects 10 and 13.

DISCUSSION

The main intent of this study was to investigate contextual effects on the representations to which a word gives rise and to determine whether or not observed variations in meaning could be given a consistent numerical representation. The results indicate that redness judgments generally do satisfy the interval order axiom. This fact is consistent with the hypothesis that context places bounds on the internal representations of a concept. The data definitely could not be described by a full order. This means that the meaning of red in each context could not be represented as a point. Instead, without exception, redness in a context had to be represented as a real interval. Since the widths of the intervals, U − L, generally varied, the overall fit of a semiorder was poor.

While acknowledging and handling homographs, most current theories of semantic memory and language comprehension (e.g., Anderson & Bower, 1973; Smith, Shoben, & Rips, 1974) have yet to deal with the fact that words systematically shift meanings as a function of context. They have not taken account of the fact that people have range or distributional knowledge. Rather, it is usually assumed that concepts can be represented as a set of features, each of which takes on a single discrete value.

How could distributional information be represented in human memory? An obvious ploy is to represent the properties of concepts as intervals instead of points. A lower limit and an upper limit, or the subjective equivalent of a mean and standard deviation, might be stored. This tack has a number of difficulties, such as the incorporation of the information in new instances and deciding which dimensions to distribute the information on. A more promising solution is to abandon the totally deductive processes which characterize most recent theories. There is evidence which suggests that people may be better at dealing with particular cases than abstracted general concepts (Anderson & McGaw,

1973; Anderson & Ortony, 1975). One could accommodate the data reported here by supposing that people generate or retrieve particular instances and base their inferences on these cases. In this view, in determining whether fire-engine red is redder than sunburn red, the subject generates the palest red fire engine and the fiercest red sunburn and compares the two. Such a process could produce the patterns of behaviors observed in this study and seems more consistent with the results from other research (Kahneman & Tversky, 1973; Tversky & Kahneman, 1971, 1973) than a system based on deduction from a set of propositions.

The deficiencies of a totally deductive approach are also suggested by research on ordered series. Several investigators have shown that the speed and accuracy of decisions about the relative position of two objects in a series increase with the distance between the objects (Lovelace & Snodgrass, 1971; Moyer, 1973; Moyer & Landauer, 1967; Scholz & Potts, 1974). An attempt to explain this fact in terms of deductive reasoning from a set of propositions is awkward, to say the least. An explanation in terms of inductive inference is far simpler and seems to fit the data from many situations (Gillman, Buckley, & Theios, Note 1).

Finally, the matter of individual differences is somewhat puzzling. Rosch (1973a, b), on the basis of cross-cultural evidence, suggests that there is considerable consistency among people in color perception. However, if our data is any indication, there is little agreement even in our own culture concerning the meaning of any particular color word.

There are several ways of approaching the subject of individual differences. One way would be to represent each context as a region in multidimensional space. This would allow the application of any of several multidimensional preference models. For example, Carroll (1972) describes a family of unfolding models relating preference to distance from some ideal. According to this theory, U and L could be represented as maximal and minimal distances from an ideal red. Within this framework, individual differences could occur in the location of the ideal red and, even more likely, in the weights that each dimension receives in the computation of distance. Another approach is suggested by the work of Tucker and Messick (1963). Each meaning could be represented by the projection of its context region on a redness vector running through the space. Individual differences would be captured in this model in the location of the vectors. Naturally, a detailed investigation of any of these models is beyond the scope of this paper.

Another possible explanation for individual differences is that, as mentioned above, the subjects might be dealing with stored distributional information instead of plausibility bounds. In other words, subjects could be dealing with the same subjective distribution of redness, but apply different criteria to those distributions to produce L and U. The low proportion of times

when one subject responded to aRb while another responded bRa is consistent with this hypothesis.

This hypothesis can be elaborated under the assumption that subjects' redness judgments involve generating particular examples. Take the sentence "The knife blade was red from the blood of the victim," for instance. One subject might construct a scenario on the assumption that the blood was fresh; for him, then, blood red would probably be interpreted as very red and comprising a narrow interval. Another subject might suppose that a detective had come upon the knife some hours after an assault; a low degree of redness would probably be inferred. Still another person might consider both scenarios, resulting perhaps in a redness interval spanning those of numerous other contexts.

The implication of the present research is that concepts are going to have to be represented in terms of context-sensitive intervals rather than context-invariant points. Several suggestions for how this might be done have been considered, but it would be premature to attempt to choose among them.

REFERENCE NOTE

1. Gillman, C. B., Buckley, P., & Theios, J. M. *Numeric comparison*. Paper presented at the meeting of the Psychonomic Society, St. Louis, November 1973.

REFERENCES

ANDERSON, J. R., & BOWER, G. H. *Human associative memory*. Washington, D.C: Winston & Sons, 1973.

ANDERSON, R. C., & MCGAW, B. On the representation of meanings of general terms. *Journal of Experimental Psychology*, 1973, **101**, 301-306.

ANDERSON, R. C., & ORTONY, A. On putting apples into bottles: A problem of polysemy. *Cognitive Psychology*, 1975, **7**, 167-180.

BARCLAY, J. R., BRANSFORD, J. D., FRANKS, J. J. MCCARRELL, N. S., & NITSCH, K. Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior*, 1974, **13**, 471-481.

CARROLL, J. D. Individual differences and multidimensional scaling. In R. N. Shepard, A. K. Romney, & S. Nerlove (Eds.), *Multidimensional scaling: Theory and applications in the behavioral sciences* (Vol. 1). New York: Seminar Press, 1972.

FISHBURN, P. C. Intransitive indifference with unequal indifference intervals. *Journal of Mathematical Psychology*, 1970, **7**, 144-149.

FISHBURN, P. C. Interval representations for interval orders and semiorders. *Journal of Mathematical Psychology*, 1973, **10**, 91-105.

KAHNEMAN, D., & TVERSKY, A. On the psychology of prediction. *Psychological Review*, 1973, **80**, 237-251.

KATZ, J. J. *The philosophy of language*. New York: Harper & Row, 1966.

KATZ, J. J. *Semantic theory*. New York: Harper & Row, 1972.

LABOV, W. The boundaries of words and their meanings. In C. J. Bailey & R. Shuy (Eds.), *New ways of analyzing variation in English*. Washington, D. C: Georgetown University Press, 1973.

LAKOFF, G. Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Papers from the Eighth Regional Meeting, Chicago Linguistics Society*. Chicago: University of Chicago Linguistics Department, 1972.

LOVELACE, E. A., & SNODGRASS, R. P. Decision times for alphabetic order of letter pairs. *Journal of Experimental Psychology*, 1971, **88**, 258-264.

MOYER, R. S. Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception & Psychophysics*, 1973, **13**, 180-184.

MOYER, R. S., & LANDAUER, T. K. Time required for judgements of numerical inequality. *Nature*, 1967, **215**, 1519-1520.

ROSCH, E. H. Natural categories. *Cognitive Psychology*, 1973, **4**, 328-350. (a)

ROSCH, E. H. On the internal structure of perceptual and semantic categories. In T. M. Moore (Ed.), *Cognitive development and the acquisition of language*. New York: Academic Press, 1973. (b)

SCHOLZ, K. W., & POTTS, G. Cognitive processing of linear orderings. *Journal of Experimental Psychology*, 1974, **102**, 323-326.

SMITH, E. E., SHOBEN, E., & RIPS, L. J. Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 1974, **81**, 214-241.

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.

TUCKER, L. R., & MESSICK, S. An individual difference model for multidimensional scaling. *Psychometrika*, 1963, **28**, 333-367.

TVERSKY, A., & KAHNEMAN, D. Belief in the law of small numbers. *Psychological Bulletin*, 1971, **76**, 105-110.

TVERSKY, A., & KAHNEMAN, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 1973, **5**, 207-231.

WEINREICH, U. Explorations in semantic theory. In D. D. Steinberg & L. A. Jakobovits (Eds.), *Semantics*. Cambridge, England: Cambridge University Press, 1971.

WITTGENSTEIN, L. *Philosophical investigations* (Trans. by G. E. Anscombe). Oxford, England: Blackwell, 1963.

NOTE

1. We have assumed strict inequality here to insure that $U > L$.