Relational judgments with remembered stimuli

DONALD G. JAMIESON and WILLIAM M. PETRUSIC
Carleton University, Ottawa, Ontario, Canada K1S 5B6

Two experiments investigating the effect of the direction of a relational judgment on the speed of the judgment are reported. In both experiments, college students required more time to select the smaller of a pair of large animals than to select the larger. Conversely, the smaller of a pair of small animals was selected more quickly than was the larger. The magnitude of this "cross-over effect" was fully graded, increasing regularly with extremity, but the variability of the response times in each direction was unrelated to extremity. Individual animals were classified as "small" or "large" with almost perfect consistency. This pattern of results is used to evaluate several models of relational judgment; of these, the congruency model is shown to be inconsistent with these data.

The time required for subjects to make a comparative judgment reflects the relation between the direction of the judgment and the position of the to-be-compared stimuli on the judged dimension. Shipley, Coffin, and Hadsell (1945), for example, observed that more time was required to select the more preferred of a pair of colors when the pair was relatively nonpreferred than when the pair was highly preferred. Shipley, Norris, and Roberts (1946) showed that the opposite effect obtained when subjects were required to select the less preferred color: now highly preferred colors were associated with slower responding than were nonpreferred colors. The interaction between the location of the judged pair on the underlying dimension and the direction of judgment, which the papers, taken together, demonstrate, has been named the "cross-over effect" (COE) by Audley and Wallis (1964). Subsequently, several papers have appeared demonstrating the COE in judgments of brightness (Audley & Wallis, 1964), pitch (Wallis & Audley, 1964), age of people in photographs (Ellis, 1972), and likelihood of the occurrence of various events (Marks[1], 1972). Thus, it seems that the COE may be a quite general property of relative judgment.

As a potential general property of relative judgment, the COE merits study for a variety of reasons. Among these, of course, is the study of the COE as an independent phenomenon—as demonstrating a yet poorly understood linguistic control over apparently straightforward comparative judgments. Methodologically, the COE greatly complicates attempts (e.g., Petrusic & Jamieson, 1975) to use

response time to measure (scale) the underlying stimulus structure when, for any of a variety of reasons, response probabilities cannot be obtained: In a pairwise scaling procedure, the effects of difficulty or of similarity, for example, would be masked by part of the interaction between stimulus location and direction of judgment that is the COE. Previous studies have viewed the COE as demonstrating a general principle of linguistic theory (Marks, 1972), as a means to a better understanding of processes underlying choice behavior, and as providing new tests of existing models of that choice process (Audley & Wallis, 1964). To determine precisely when each of the above motivations for the study of the COE may be relevant requires a much greater understanding of the properties of the COE. At present, few COE properties, beyond occurrence, are established. For example, the effects of changes in task difficulty or the speed-accuracy tradeoff upon the COE have not been systematically investigated. In fact, the conditions under which the COE occurs are not well distinguished from those in which a weaker condition—the funnel effect—holds (the funnel effect is said to occur when response latencies differ for judgments of pairs at one end of the dimension but are comparable for judgments at the other end). The present paper examines one property—the differential effects of direction of judgment upon response time for stimuli *distributed* along the dimension (here, size). Three simple cross-over effects are possible (Wallis & Audley, 1964). First, the effect may occur to a greater or lesser extent along the entire dimension (i.e., the COE may be graded). Second, the effect may occur only for stimuli near the extremes of the dimension. Finally, the effect may occur to the same degree for all stimuli, but in an opposite direction for small and large stimuli. Several of the models proposed to account for the COE make clear predictions concerning this gradedness property. These models are discussed below. As well, some views of relational judgment treat the COE as being very specific. For example, in the system outlined by

Clark, Carpenter, and Just (1973), or according to the lexical markedness model proposed by Marks (1972), at most a funnel effect should be observed with size judgments.

Chronologically, the first of the various explanations offered for the COE was the extension of Audley's (1960) stochastic choice model to the COE observed by Audley and Wallis (1964) in brightness comparisons. In this application, the final, overt response is made only after a number of implicit, covert responses have occurred. That is, the subject responds only after accumulating (observing) a criterion run[2] of one of the types of implicit, covert response. In the model applied to the Audley and Wallis experiments, two implicit responses—"bright" and "dark"—occurred in a random order, but at *rates* determined by the brightness of the stimuli. Thus a bright stimulus produced a high rate of implicit "bright" responses and a low rate of implicit "dark" responses, while a dark stimulus produced a low rate of "bright" responses and a high rate of "dark" responses. The Audley and Wallis model viewed the choice process as a type of race between the implicit responses to the two stimuli of a pair. If the instruction was "choose brighter," the subject selected the first stimulus to be associated with a criterion run of "bright" implicit responses. If the instruction was "choose darker," the subject selected the first stimulus to be associated with a criterion run of "dark" implicit responses. Since, thus, the instruction determined which implicit response would be monitored on a trial, two bright stimuli presented with the instruction "choose brighter" would likely result in the rapid accumulation of the required run of implicit "bright" responses, and overt responding would be relatively fast. When required to choose the darker of two bright stimuli, however, the overt response would be delayed since more time would be required for a criterion run of "dark" responses to occur. The application of the reverse argument when dark stimuli are presented leads to the prediction of the COE. Note that a graded COE occurs as a natural consequence of the Audley and Wallis model, since the rates at which the implicit responses are emitted depend upon the location of the stimuli on the judged dimension.

In a second model, proposed by Clark and his associates (Banks, Clark, & Lucy, 1975; Clark, Carpenter, & Just, 1973), Marks (1972), and Wallis and Audley (1964), and named the congruency model, following Marks, the COE evolves in a quite different way. In this case, the initial overt coding of a stimulus pair is determined, not by the instruction, but by the stimuli. Hence, two bright lights are coded as bright—i.e., in the form "bright," "bright +," while two dark lights are coded in the form "dark," "dark +" (the + represents the more extreme, e.g., darker, of the pair). The overt response time, however, depends on the *relation* between the

stimulus code and the instruction. When the required direction of judgment does not match the coded stimulus presentation, a translation is required, delaying the response. This translation will be required whenever subjects select the brighter of two stimuli which are coded as "dark," and when they select the darker of two stimuli which are coded as "bright." Thus, in those cases, response times will be inflated relative to the times to select the brighter of two "bright" or the darker of two "dark" stimuli—i.e., relative to "congruent" judgments. Note that the congruency model would require any grading of the COE to occur in the *coding* of the stimuli (e.g., as bright or dark), i.e., the probability that a particular stimulus is coded as bright/dark must increase/decrease with increasing luminance. If this coding is not probabilistic, the COE must be of a constant size for all "large" and for all "small" stimuli, i.e., the third type of simple COE grading must apply.

Interestingly, however, the system of Clark and his associates seems not to admit the COE at all for some comparisons, notably for *size* comparisons, because of linguistic assumptions. Specifically, Banks et al. (1975, p. 38) and Clark et al. (1973, pp. 340-350) suggest a single underlying dimension for size—"larger" but not "smaller." Thus stimuli which differ in size should be codable only in terms of "large" and "large +," not in terms of "small" and "small +." Hence, the translations which generate the COE in this model can only occur for the instruction "choose smaller." In all cases, therefore, "choose larger" responses should be faster than "choose smaller" responses, and no COE would be observed.

A third model, the discriminal dispersion model, was proposed by Marks (1972) as an explanation of the COE. In this model, the subject assigns an ideal point which depends on the instruction (e.g., an ideal large when the instruction is choose larger), against which stimuli are compared. Response time in this model is inversely related to the discriminability of the two stimuli and directly related to the variances of the stimuli. To make response times instruction-dependent, yielding the COE, Marks required the variances of the stimulus representations to increase with the distance of a stimulus from the relevant ideal point. Grading appears as a natural consequence of Marks' model, but the dependence of the actual *representation* of a stimulus on linguistic stimuli—the instruction—may be unnecessarily strong.

The experiments reported here sought to establish the occurrence of the COE in size comparisons and to examine the gradedness property of the COE. In particular, the demonstration of grading of the COE would permit a detailed examination of the alternative models—in which grading results from regular changes in the *proportion of* trials on which *translations* must occur (the congruency models of

Banks et al., Clark et al., Marks, and Wallis & Audley), from changes in the *rates* at which the relevant predecisional information is acquired (Audley & Wallis), or, with a suitable choice of latency function, from changes in the *representation of the stimuli* under the different instruction conditions (Marks' discriminal dispersion model). Because of the convenience of verbal materials, and because of the intrinsic interest of internal comparisons, animal names from Moyer's (1973) pool were selected as stimuli.

METHOD

Experiment 1

Thirty-two introductory psychology students served for approximately 15 min in return for course credit. Two sets of stimuli were drawn from Moyer's (1973) item pool. The first set contained six five-letter animal names (louse, quail, sheep, horse, moose, whale), while the second set contained six three-letter animal names (ant, bee, rat, cat, hog, cow). Instructions [(choose) smaller or larger] and stimuli were written horizontally, in letters .63 cm high and .28 cm wide, near the center of an Electrohome video monitor. Timing, sequencing, presentation of instructions and stimuli, and recording of responses and response times were controlled by a DataGen Nova 1220 computer. Each of the animal names also appeared on a 3 x 5 in. index card.

Each trial consisted of the following sequence of events: a 2-sec delay; the presentation of the instruction—a single word ("larger" or "smaller")—for 750 msec; a 500-msec delay; the presentation of a pair of animal names (in a horizontal line, separated by two spaces) which remained on the screen until the subject responded; a delay while the subject rested before pressing a bar to initiate the next trial.

Each subject received 120 trials with one of the two stimulus sets—three- or five-letter animal names. Within each of four blocks of 30 trials, each of the 30 pairwise combinations of the six names within a group was presented exactly once, with presentation order determined randomly. Instructions remained unchanged within a block but changed between blocks. For eight subjects in each group the instruction order was 30 trials choose larger, 60 trials choose smaller, 30 trials choose larger (30L, 60S, 30L). For the remaining eight subjects in each group, the order 30S, 60L, 30S was used.

The subject was seated in front of a table which supported the video monitor and a panel containing two response keys. The subject was instructed to press the key to the side of the correct animal and to respond as quickly as possible without making errors. Before the trials were presented, the subject was given the six cards with the animal names for his group, and told that these would be the animal names he would be comparing. He was asked to read each of the names and then to place the cards in the order of animal size.

Experiment 2

The stimulus set from Experiment 1 was modified by adding the word "elk" to the three-letter set and the word "roach" to the five-letter set. A four-letter set (flea, moth, dove, frog, duck, wolf, bear) was also included. Two experimentally naive undergraduate females were paid an hourly rate to participate. Subject 1 served for six sessions with the four-letter set, followed by six sessions with the five-letter set and then six sessions with the three-letter set. Subject 2 served for six sessions with the three-letter set, followed by six sessions with the five-letter set. During each of the six sessions with a particular set, a subject made 192 judgments—eight under each instruction condition with each of the 12 (ordered) pairwise adjacent combinations of the seven names within the set. Each session thus consisted of 16 blocks, each of 12 trials. During a

block, each of the 12 pairwise adjacent combinations of the stimuli within a set was presented exactly once, with presentation order determined randomly. Instructions remained constant within a block of 12 trials but alternated between blocks. Subjects rested for 5 min after every fourth block. Other aspects of the procedure were as described for Experiment 1.

At the completion of the comparison sessions, each subject served for a further session, during which she classified each of the animals as small or large. On each classification trial, the word "ready" was presented for 500 msec, followed, after a 500-msec delay, by one of the animal names. This name remained on the video screen until the subject pressed one of two preassigned buttons to indicate that she considered that animal to be "small" or "large." A 1-sec intertrial interval followed each response. During the session, each subject made 20 classifications with each of the animals within her comparison sets. These classifications were arranged in 10 blocks of 14 trials each, with each block containing two presentations of each of the seven names within a set, in a completely randomized order. All classifications with one set were completed before the next set was presented.

RESULTS

The COE was examined primarily with judgments on animal pairs adjacent in the physical size ordering. On approximately 1.46% of trials in Experiment 1 and on 1.65% of trials in Experiment 2, incorrect responses occurred. These incorrect responses were discarded prior to any analyses. Figure 1, which presents, for Experiment 1, the mean latencies to compare the extreme adjacent pairs under the two instruction conditions, shows that the full COE
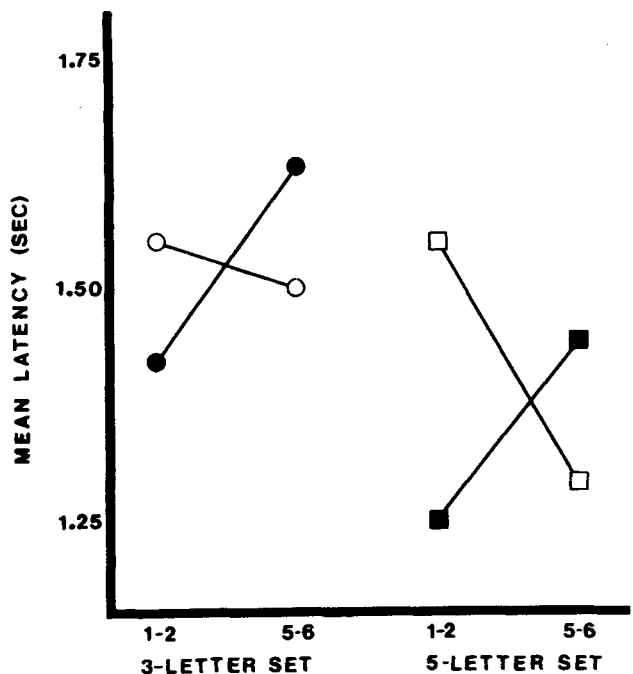

Figure 1. The cross-over effect in Experiment 1. Time (in seconds) to select the larger animal (open points) and the smaller animal (closed points) for the five-letter set (squares; pairs "louse quail" and "moose whale ") and for the three-letter set (circles; pairs "ant bee" and "hog cow").
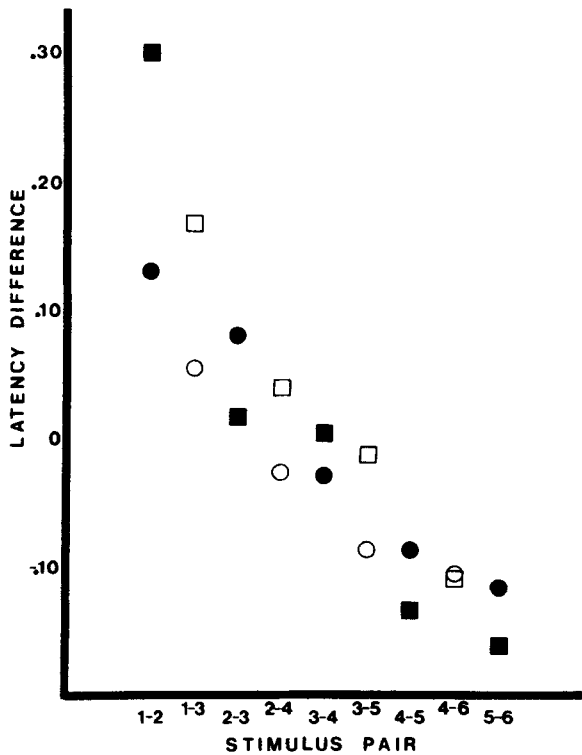
Figure 2. The difference (time to select larger animal — time to select smaller animal) for the adjacent pairs (filled points) and pairs separated by two steps (open points) from the three-letter sets (squares) and the five-letter sets (circles) of Experiment 1. Each point is the mean of 16 difference scores—those computed from the mean response times of each subject within a group on a given pair. Difference scores are in seconds.

obtained for both the three- and the five-letter animal sets: The larger of the two large animals was selected more rapidly than was the smaller, while the larger of the two small animals was selected more slowly than was the smaller. Thus, the particular assumption of the Clark et al. model may be incorrect, since through it the congruency model predicts no COE in the present situation.

As the closed points in Figure 2 show, plotting, for adjacent pairs, the mean across subjects of the difference (time to select larger - time to select smaller), yields a finely graded plot which is 125-300 msec in favor of judgments of "smaller" for the smallest animals (Pairs 1-2) and 130-150 msec in favor of "large" judgments for the largest animals (Pairs 5-6). Since, for each of the two animal sets represented in Figure 2, there are 5! = 80 possible orderings of the latency differences, the probability of obtaining by chance, twice, the observed grading is $(1/80)^2 < .0002$. Also included in Figure 2 (open points) are the mean differences of judgment times for stimulus pairs separated by two steps (e.g., the pair 1-3). These points fit rather well on the adjacent-points plot, emphasizing the gradedness of this COE. Further, a comparison of the slopes of the

adjacent points and the two-step plots permits an evaluation of the effects of difficulty on the magnitude of the COE. If the COE magnitude increased with the difficulty of the judgment, the slopes of the adjacent and one-step curves would differ; the adjacent-pairs curve would have a steeper slope than would the one-step curve. In fact, the adjacent-points slope is larger for both stimulus sets when straight lines are fitted by eye. Banks et al. (1975) suggested that the COE magnitude is independent of judgment difficulty; the present results suggest some caution on this point. Moreover, since the difficulty manipulation, both here and in the Banks et al. experiment, was supraliminal, i.e., even the most difficult comparisons were essentially error-free, we feel that a conclusion on the dependence of COE magnitude on task difficulty should await a more complete study of difficulty—one in which a wider range of difficulty, including difficulty measured in terms of response probability, is examined.

Since the observed grading of Figure 2 might be attributable to averaging data over subjects, each of whom demonstrated an all-or-none COE but with the COE reversing direction at different stimulus pairs, the individual subject data of Experiment 2 was obtained. To use the entire distributions available in Experiment 2 to examine the effects of instruction on the response times for each pair, an estimate of the probability that a randomly selected "choose larger" response time would exceed a randomly selected "choose smaller" response time was calculated. This estimate was obtained for each pair by plotting the proportion of the "select smaller" response times which did not exceed each point of the cumulative distribution of the "select larger" response times, and then calculating the proportion of area falling under this curve. These proportions, which are presented in Figure 3, show that the COE was also clearly graded for the individual subject data from Experiment 2. For Subject 1, the estimated probabilities decrease monotonically with the physical size of the animal pairs for both the three- and the four-letter sets; the five-letter set shows a single pairwise violation of monotonicity. Subject 2, a bilingual francophone, shows a single pairwise violation of monotonicity for the three-letter animal names; with the five-letter stimuli, the trend is disrupted for Pairs 3 and 4. After averaging over the different sets for each subject, the differences are perfectly monotone for Subject 1, while Pair 3 remains a single violation of monotonicity for Subject 2. On the basis of these data, therefore, the hypothesis that the graded COE found in Experiment 1 is attributable to averaging data over subjects cannot be supported.

As noted, the congruency model predicts grading only as a consequence of probabilistic coding of the intermediate stimuli. The classification data, presented in Table 1, however, show that both subjects were almost perfectly consistent in their
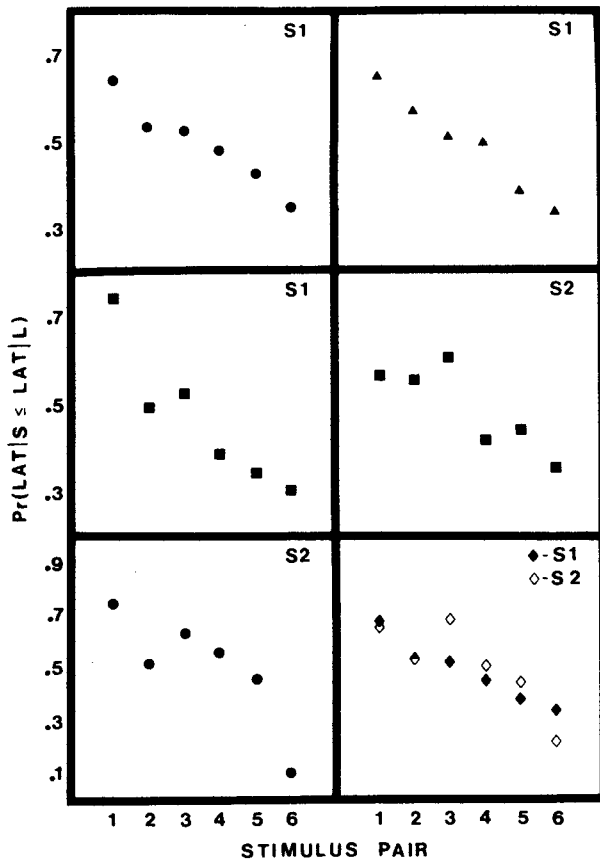
Figure 3. Estimates of the probability of a "smaller" response occurring at least as quickly as a "choose larger" response for each subject for the three-letter (circles), four-letter (triangles), or five-letter (squares) pairs. Each estimate was obtained by calculating the proportion of area falling under the plot of the proportion of "choose smaller" response times which were at least as fast as each point of the observed cumulative distribution of "choose larger" response times. Diamonds are the mean proportions (across stimulus sets) at each point in the within-set size orderings, for Subject 1 (closed points) and for Subject 2 (open points).

classification of each pair. To the extent, then, that these classifications represent the codings of the stimuli, the probabilistic coding notion is disconfirmed. As well, if probabilistic coding does occur, then, within each instruction condition, response times must be less variable for extreme stimulus pairs (where translations occur on most trials under one instruction and on few trials under the opposite instruction) than for intermediate pairs (where numerous translations occur under each instruction condition). The standard deviations of the response times for each pair, computed as a test of probabilistic coding, are presented in Table 2. Each entry in this table represents the arithmetic mean of four standard deviations computed separately for each pair of stimuli under the two instructions and two presentation order (ant-bee or bee-ant) conditions. As Table 2 shows, standard deviations do

not differ systematically between extreme and intermediate pairs. Thus we find no support for the postdecision probabilistic response translation hypothesis which must underlie the prediction of a graded COE by the congruency model.

GENERAL DISCUSSION

Moyer (1973) has demonstrated that unidirectional size comparisons with these animal-name stimuli have the properties that error rates and response times are inverse functions of the physical size difference of the two animals compared. The present demonstration of the COE in the comparison of remembered animals encourages and extends Moyer's analogy of an internal psychophysics. Moreover, barring any unique factors associated with the present task, the occurrence of the full COE in size comparisons with these stimuli suggests that the Clark et al. assumption of a single dimension (largeness) for size judgments is in error.

The present demonstration, that the COE, in the data of individual subjects, is graded and not all-or-none, is a necessary condition for the verification of both the stochastic model of Audley and Wallis (1964) and the discriminal dispersion model of Marks (1972). The congruency model of Banks, Clark, and Lucy (1975), Clark, Carpenter, and Just (1973), Marks (1972), and Wallis and Audley (1964) could also predict a graded COE, but the required probabilistic coding is contraindicated by the classification data of Experiment 2 and by the pattern

Table 1
Number of Times (Out of 20) that Each Subject Classified a Stimulus as "Small"

Subject	Stimulus Set	1	2	3	4	5	6	7
1	3s	20	20	20	20	3	0	1
	4s	20	20	20	20	20	0	0
	5s	20	20	20	0	0	0	0
2	3s	20	20	19	20	3	0	1
	5s	20	20	20	9	0	1	0

Table 2
Mean Standard Deviation of Response Latencies for Each Pair of Stimuli from Experiment II

Subject	Stimulus Set	1	2	3	4	5	6
1	3s	223	260	248	176	301	175
	4s	283	236	276	237	229	205
	5s	220	190	208	186	189	197
2	3s	333	355	303	334	334	298
	5s	333	304	326	317	293	290

Note—Entries in milliseconds.

of variances of response times for the various pairs. Thus, only the dispersion model of Marks and the Audley and Wallis model are supported by the present data. An interpretive problem with the dependence of the *noise of the stimulus representation* upon the linguistic instruction remains, however, for the Marks model. A model which is similar in nature to Marks' but which does not make this strong, intuitively and epistemologically troublesome assumption is the *discrepancy ratio model*. This version of a model, considered in a preferential choice situation by Greenberg (1963), retains the unfolding theory (Coombs, 1950) notion of an ideal point associated with the direction of the comparison. However, Marks' distributional assumptions are discarded. The discrepancy ratio model is discussed for the unidimensional case, below.

Let x and y be stimuli with locations s(x) and s(y), respectively, on the unidimensional attribute under consideration (size). Let I^O represent the ideal point for "large" and let I_O represent the ideal point for "small." It is assumed that subjects select the larger of x and y by comparing the distance, denoted $d[I^O,s(x)]$, of the stimulus x from the "large" ideal point, with the distance $d[I^O,s(y)]$, of stimulus y from the ideal point according to the following decision rule. Let $R = d[I^O,s(x)]/d[I^O,s(y)]$ denote the discrepancy ratio, and C^O a positive criterion constant (equal to one in the unbiased case as in Greenberg's, 1963, model) such that:

x is larger than y if and only if $R \leqslant C^O$

y is larger than x if and only if $R > C^O$.

The time, T, required for the judgment to be completed is assumed to depend on the distance R is from its criterion, C, denoted D(R,C), so that

$$T = f[D(R,C)] + A$$

where A is a positive constant not involving decision processing and f is some suitable monotone decreasing function. The model in this form has the desirable properties that the stimulus representation is independent of the direction of judgment; the response times are dependent on the direction of judgment, being relatively more rapid for the "congruent" judgment than for the "incongruent" judgment, since the discrepancy ratio is larger in the latter case; the COE is symmetrically graded about the bisector of the two ideal points; manipulation of the criterion (bias) parameters C^O and C_O provides a convenient way for various forms of the COE, including the funnel effect, to be generated; and response times decrease with increasing distance between the stimuli (since then, relative to either ideal point, ratios decrease). Of course, specification of the

function f, perhaps following Thomas (1971), permits more precise predictions to be made.

REFERENCES

AUDLEY, R. J. A stochastic model for individual choice behaviour. *Psychological Review,* 1960, **67**, 1-15.

AUDLEY, R. J., & PIKE, A. R. Some alternative stochastic models of choice. *British Journal of Mathematical and Statistical Psychology,* 1965, **18**, 207-255.

AUDLEY, R. J., & WALLIS, C. P. Response instructions and the speed of relative judgments. I. Some experiments on brightness discrimination. *British Journal of Psychology,* 1964, **55**, 59-73.

BANKS, W. P., CLARK, H. H., & LUCY, P. The locus of the semantic congruity effect in comparative judgments. *Journal Experimental Psychology: Human Perception and Performance,* 1975, **1**, 35-47.

CLARK, H. H., CARPENTER, P. A., & JUST, M. A. On the meeting of semantics and perception. In W. G. Chase (Ed.), *Visual information processing.* New York: Academic Press, 1973.

COOMBS, C. H. Psychological scaling without a unit of measurement. *Psychological Review,* 1950, **57**, 145-158.

ELLIS, S. H. Interaction of encoding and retrieval in relative age judgments: An extension of the "cross-over" effect. *Journal of Experimental Psychology,* 1972, **94**, 291-294.

GREENBERG, M. G. J-scale models for preferential choice behavior. *Psychometrika,* 1963, **28**, 265-271.

MARKS, D. F. Relative judgment: A phenomenon and a theory. *Perception & Psychophysics,* 1972, **11**, 156-160.

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

SHIPLEY, W. C., COFFIN, J. L., & HADSELL, K. C. Reaction time in judgment of color preference. *Journal of Experimental Psychology,* 1945, **35**, 206-215.

SHIPLEY, W. C., NORRIS, E. D., & ROBERTS, M. L. The effect of changed polarity of set on decision time of affective judgments. *Journal of Experimental Psychology,* 1946, **36**, 237-243.

THOMAS, E. A. C. Sufficient conditions for monotone hazard rate: An application to latency-probability curves. *Journal of Mathematical Psychology,* 1971, **8**, 303-332.

WALLIS, C. P., & AUDLEY, R. J. Response instructions and the speed of relative judgments. II. Pitch discrimination. *British Journal of Psychology,* 1964, **55**, 133-142.

NOTES

1. In fact, as Marks notes, with likelihood judgments, funnel effects were more frequent than were cross-over effects.

2. Of course, other stochastic counter models (see Audley & Pike, 1965) are readily substituted here since both the COE and gradedness follow from the assumptions of the implicit response distribution.

3. Gradedness is in no way peculiar to this measure. Plots of median or mean latency differences show similar grading.