Memory & Cognition
1992, 20 (3), 254-270

The discriminability of remembered magnitudes

JOSEPH V. BARANSKI and WILLIAM M. PETRUSIC
Carleton University, Ottawa, Ontario, Canada

Psychometric functions and the associated indices of discriminative performance (i.e., the point of subjective equality [PSE], just noticeable difference, and Weber fraction) were obtained with the method of constant stimuli using *perceptual* and *remembered* line-length standards. Three important results were obtained. First, comparisons with a perceptual or a remembered standard were sensitive to variations of absolute stimulus differences with a common ratio; that is, Weber's law was violated. Second, relative to discriminative performance with the longest and shortest remembered standards, comparisons involving mid-range remembered standards displayed increased variability in the PSE and inflated Weber fractions, characteristic of a reduction in the *quality* of the memorial representation. Finally, large and negative time-order errors (TOE) were observed for successive line judgments but not for those involving remembered standards. The implications of these findings for research concerned with the relationships between perception and memory, as well as the TOE phenomenon, are discussed.

Woodworth (1938), in the first chapter of his classic text, "Experimental Psychology," concludes:

> The experimental results are in good agreement on the negative conclusion: No absolute difference exists between an image and a percept, and there is no sure criterion by which one can be distinguished from the other. (p. 45)

Following the recent resurgence of interest in the relationships between perception and memory, Woodworth's (1938) negative conclusion can be viewed more positively. Several findings, based on the application of traditional psychophysical methods to remembered stimuli, have revealed important commonalities between perception and memory, and these commonalities permit a characterization of their shared processes and properties (see Finke, 1980, 1985; Finke & Shepard, 1985; Kosslyn, 1980; Shepard, 1984, for reviews).

Shepard and Chipman (1970) originally provided the elegant demonstration of a "second-order isomorphism" between perception and memory, derived from a multidimensional scaling analysis of perceived and remembered representations of the shapes of the U.S. states. Similarly, several experiments using magnitude estimates (Stevens, 1957) of perceived and remembered stimuli have demonstrated an important psychophysical relationship on various continua (e.g., Bjorkman, Lundberg, & Tarnblom, 1960; Bradley & Vido, 1985; Chew & Richardson, 1980; Kerst & Howard, 1978; Moyer, Bradley, Sorensen, Whiting, & Mansfield, 1978). In each case, the perceived and remembered magnitude estimates were power functions of the physical magnitudes with related exponents. As well, Algom, Wolf, and Bergman (1985) and Wolf and Algom (1987) have demonstrated similar commonalities between perception and memory on the basis of psychophysical relations (power functions) derived from functional measurement analyses (Anderson, 1970).

In the context of psychophysical comparison, it has long been known that as two perceived magnitudes become more similar for comparison, more time is required to distinguish them (Cattell, 1902; Henmon, 1906; Münsterberg, 1894). Moyer (1973; Moyer & Landauer, 1967) demonstrated the important parallel for memory by showing that a response time *distance effect* can be obtained for comparisons involving subjective differences between symbolically coded items. Soon after, Moyer and Bayer (1976) extended the symbolic comparison research to include magnitudes compared on an induced linear ordering (cf. Phillips, 1958). They first employed a paired-associate learning phase where four different circle sizes were learned to be represented by specific consonant-vowel-consonant (CVC) letter triads. In the subsequent comparative judgment task, subjects judged the relative size of either pairs of circles or pairs of CVCs. In addition to distance effects, Moyer and Bayer (1976) found that the stimulus set with larger absolute differences between items (i.e., larger stimulus range) displayed faster response times for both the perceived and remembered comparisons. This *range effect* led them to conclude that an interval scale representation of perceptual magnitudes could be preserved in memory and that subjects can use this interval (analogue) information when performing mental comparisons.

Subsequently, Banks, Mermelstein, and Yu (1982) and Henderson and Well (1985) failed to replicate the Moyer and Bayer (1976) memory range effect; suggesting that, at best, an ordinal scale relation holds for remembered magnitudes. However, perhaps less well publicized was

This research was supported by a grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada to W.M.P. We thank Robert Proctor and three anonymous reviewers for their helpful comments. Correspondence can be addressed to either author at the Department of Psychology, Carleton University, Colonel By Drive, Ottawa, ON, Canada K1S 5B6.

the fact that Banks et al. (1982) also did not find a reliable range effect for the perceptual comparisons. Henderson and Well (1985), on the other hand, did not provide a perceptual condition in order to contrast the remembered comparison data. It may be, then, that the strongest evidence against the analogue view of memorial representation stemmed from instances in which the stimulus range was not sufficiently different for memory or perception (see Petrusic & Baranski, 1992, for a reexamination of the range effect).

The present studies continue to explore the commonalities and differences between perception and memory in the context of relative judgments by extending the classic psychophysical method of constant stimuli (Hegelmaier, 1852) to include remembered standards. Thus, on all trials, subjects chose the longer or the shorter of two horizontal line segments, where one was a physical, variable, "comparison" stimulus. On one half of the trials the "standard" was also a physical line, whereas on the other half the standard was a letter triad (CVC) that represented the physical standard (defined, as with Moyer & Bayer, 1976, in a previous paired-associate learning session).

With the method of constant stimuli, psychometric functions can be obtained separately for comparisons with the perceptual and the remembered standards. Psychometric-function analyses permit a direct comparison of discriminative sensitivity with a perceptual and a remembered standard, characterized in terms of the classic measures of just noticeable difference (JND) and Weber fraction. In addition, psychometric-function analyses also permit derivation of a point of subjective equality (PSE) for a perceptual and a remembered standard. Importantly, the difference between the PSE for a remembered standard and the actual, physical, standard value (i.e., the *constant error*) provides an index of the location precision for remembered magnitudes.

The task employed *successive* comparisons because Ono (1967) has shown that neither the strict form of Weber's law nor the generalized Weber's law proposed by Miller (1947) (JND $= WF \times St + c$) holds for simultaneous line comparisons because subjects can solve the discrimination problem by simply noting discrepancies at the ends of the two lines (i.e., the task becomes tantamount to an acuity judgment based on stimulus differences). Importantly, Ono also showed that this constraint can be overcome when the two stimuli are spatially separated and thus not available in the same visual fixation. However, although Ono's latter result provides a sufficient condition for the comparison of simultaneously presented lines, spatially separated presentation of a line and a CVC still permits possible identification/categorization of the comparison lines relative to the CVC. Hence, only by presenting the stimuli successively can we ensure *comparison* on trials involving line standards and CVC standards. Of course, when two lines are presented successively, the comparison necessarily requires memory for the first line. Thus, comparisons of this variety are commonly referred to as "perceptual-memory" comparisons (see Kinchla & Smyzer, 1967; Laming & Scheiwiller, 1985). Hence, the present studies sought to extend the examination of Weber's law for linear extent to perceptual-memory comparisons and comparisons involving (long-term) symbolic representations of perceptual magnitudes.

An additional benefit of successive comparisons is that they permit the study of *presentation-order effects* with the perceptual and symbolic standards. Presentation-order effects were first discovered by Fechner (1860/1966), who found that the discrimination of a pair of successively lifted weights was better when the second weight lifted was the heavier. Fechner named this underestimation of the first stimulus a negative error and contrasted it with a positive error (typically found with small magnitudes) in which the first stimulus is overestimated relative to the second.

Since Fechner's (1860/1966) findings, a very large empirical and theoretical literature has arisen concerning these presentation-order effects, or *time-order errors* (TOEs) as they are more commonly known (see Hellström, 1985, for a comprehensive review). Although TOEs have been extensively studied with successive perceptual stimuli, especially in the context of duration discrimination (e.g., Allan & Kristofferson, 1974; Hellström, 1977; Jamieson & Petrusic, 1975a, 1975c; Petrusic, 1984; Schab & Crowder, 1988), they have yet to be investigated for judgments involving symbolic magnitudes.

The occurrence of TOEs with symbolic magnitudes would be especially interesting because several recent interpretations of the TOE consider the effect to be a perceptual-memory phenomenon (Hellström, 1977, 1985; Jamieson & Petrusic, 1975a, 1975c; Petrusic, 1984). For example, the prevalent theoretical position of the TOE to date is Hellström's (1985) quantitatively developed differential sensation-weighting theory, where the weights are understood to reflect dynamic contrast and assimilation effects operating on the perceptual-memory representations of the stimuli to be compared. Hence, according to this view, the occurrence of TOEs with symbolic magnitudes would imply that a symbolic CVC code can activate the same sensory mechanisms involved in direct perceptual experiences. Interestingly, this general position has been proposed by proponents of the strong *imagery* view of memory representations (e.g., Farah, 1985; Finke, 1980, 1985).

In summary, one goal of the research was to obtain PSEs and JNDs for long-term memories of percepts by requiring, for the first time, *confusable* relational judgments with symbolic magnitudes. The second goal was to investigate presentation-order effects with symbolic magnitudes. The occurrence of TOEs with symbolic magnitudes would extend the research that suggests a functional equivalence between perception and memory and have important implications about the nature of symbolic representations. On the other hand, a failure to obtain TOEs with symbolic magnitudes would provide a criterion for distinguishing between the processing of perceptual and memory-based information.

EXPERIMENT 1

Method

Subjects

Twenty-three Carleton University students participated in one 90-min session to satisfy part of the laboratory requirements of a third-year cognition course. All subjects reported normal or corrected-to-normal vision and were naive with respect to the nature and aims of the experiment.

Apparatus

Stimuli were presented on an Amdek-310A video monitor. High-resolution graphics were permitted with a Hercules monochrome card and MetaWindows graphics under Turbo Pascal software control. Timing, accurate to within ±1 msec, was possible with a Data Translation clock board and extensive software development. Graphics production, stimulus presentation, event sequencing, and the recording of responses and response times were controlled by an IBM PC XT clone computer. Responses were made using the buttons on an IBM PC mouse, and confidence levels were registered using the PC keyboard.

Stimuli

Three horizontal lines (10.54, 100, and 200 mm) were used as *standards* in the method of constant stimuli. Five *comparison* line lengths were chosen at −6%, −3%, 0%, +3%, and +6% difference from each standard. Thus, the comparison lines had values of 9.90, 10.22, 10.54, 10.86, and 11.18 mm for the short standard, 94, 97, 100, 103, and 106 mm for the medium-length standard, and 188, 194, 200, 206, and 212 mm for the long standard. Hence, as the length of the standard increased, the stimulus *ratios* were held constant but the absolute stimulus *differences* (and thus a priori ease of comparison) increased.

All lines were 1 mm wide and appeared in amber color on a black background. The midpoint of each line was horizontally and vertically centered on the screen, and the shortest and longest lines subtended visual angles of approximately 1° and 10°, respectively.

Procedure

Part 1: Learning. Each session began with a learning phase in which the subjects learned to associate the three line-length standards with specific letter triads (GUF, BIX, or ZOC), which were counterbalanced across subjects according to a Latin-square design. On each learning trial, a single line was presented on the computer screen, below which was a rectangular box, divided into three equal squares, in which the three possible CVCs appeared (in a random order on each trial). By moving the PC mouse from side to side, the subject was able to illuminate, in succession, any one of the three CVC-containing squares. The subject depressed the middle key on the mouse when he/she assumed the appropriately illuminated CVC to correspond to the line length presented. The subject had 5,000 msec to perform the association. If this time was exceeded, "Too Slow" appeared and the appropriate box was illuminated for the subject for 2,000 msec. If the subject responded under the time limit but was incorrect, then again the appropriate box was illuminated for 2,000 msec.

The subjects were randomly assigned to one of two learning groups. The subjects in Group 1, the "underlearning" group, performed these associations until they met a criterion of 9 successive matches, 3 with each CVC standard. This was then followed by 18 additional learning trials. The subjects in Group 2, the "overlearning" group, also had to provide 9 successive matches, but this was followed by 90 additional learning trials. All subjects were instructed to closely attend to the magnitudes of the line lengths they were learning since they would be using this information in the next part of the experiment.

Part 2: Comparative judgment task. Following the learning phase, the subjects were instructed that on each trial they would be presented, in succession, either two horizontal lines, a line and a CVC, or a CVC and a line. In each case, they would have to select, according to the instruction, either the longer or the shorter of the two presentations. In the case of either a line-CVC or CVC-line presentation, the CVC was to correspond to the line length learned in the previous phase of the experiment. For all trials, the left button on the mouse corresponded to the stimulus presented first and the right button to that presented second, relative to the instruction. The subjects were asked to be as accurate as they could, while not taking too much time to respond.

Each session was comprised of 3 blocks of 120 randomized trials. The trials in each block arose from the factorial combination of the 3 standard lengths (10.54, 100, and 200 mm) × 2 types of standard (line, CVC) × 5 comparison stimuli (−6%, −3%, 0%, +3%, +6% difference) × 2 instructions (LONGER, SHORTER) × 2 presentation orders (standard first, standard second).

Each trial began with the presentation of an instruction (LONGER or SHORTER), which remained centered near the top of the screen throughout the trial. One second after the instruction appeared, the first stimulus was presented for 2,000 msec and then removed. Following a 500-msec interstimulus interval (ISI), the second stimulus was presented and remained on the screen until the completion of the comparative judgment. Response times were recorded from the presentation of the second stimulus until response selection. Following the response, the screen was cleared and the subject was given a visual prompt to provide a confidence rating that was to indicate how certain he/she was that he/she made the correct judgment.[1] The subject typed a confidence value (0-100) on the PC keyboard and then depressed the "Enter" key to record the confidence report and initiate the next trial.

The subject operated under a 10-sec deadline. If this was exceeded, the statement "Too Slow" appeared following the response. Two seconds separated the registration of confidence and the next trial, and the subject was not provided with accuracy feedback at any point during the task.

Results and Discussion

Three subjects provided pilot sessions in order to establish the comparative difficulty levels and sessional duration. The data of these subjects are not reported. For the remaining 20 subjects, trials on which response times exceeded 10 sec were discarded. This accounted for 28 of the 7,200 trials obtained (0.3%).

The results are presented in four major sections. The first outlines the effects of the differential learning criterion. The second provides the psychometric-function analysis and associated indices of discrimination performance. The third investigates the effects of presentation order, and the fourth examines the response-time properties associated with the comparative judgment task.

Differential Learning Effects

A five-way analysis of variance (ANOVA), with (arcsine transformed) percent correct as the dependent measure, was used to analyze the results of the differential learning criterion. The two learning conditions provided the between-group factor, and standard type (line, CVC), standard length (10.54, 100, and 200 mm), variable stimulus,[2] and presentation order (long-short vs. short-long) were within-subject variables.

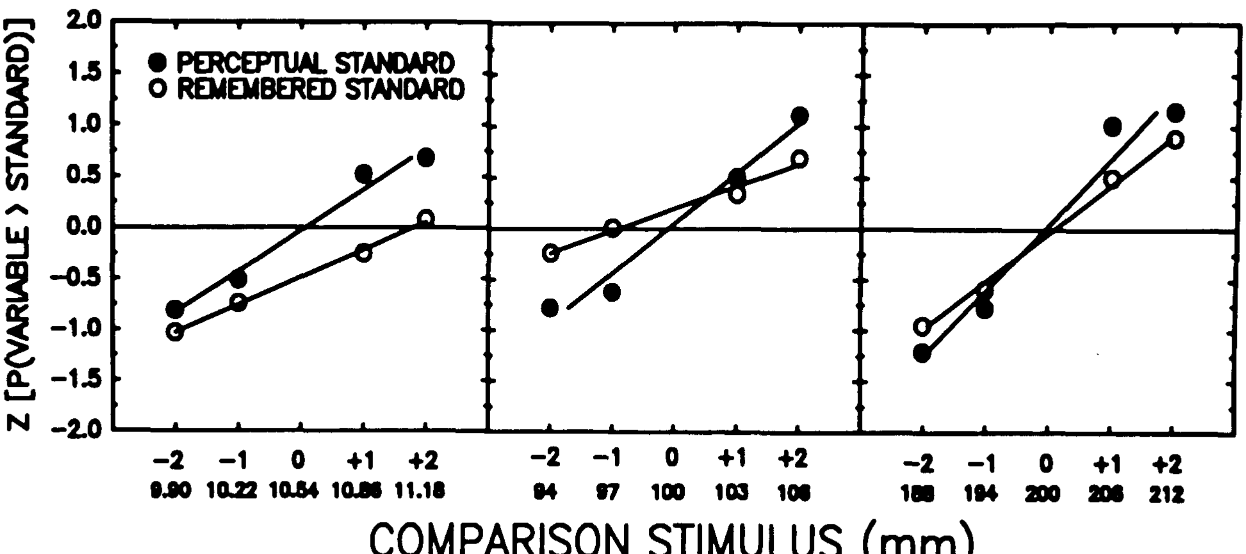

Figure 1. *Z* score psychometric functions for the three levels of perceptual (filled circles) and remembered (open circles) standards in Experiment 1. R^2 values for the 10.54-, 100-, and 200-mm standards are, respectively, .996, .990, and .993 for the remembered standards and .978, .976, and .960 for the perceptual standards.

The main effect of learning group was reliable [$F(1,18) = 7.82$, $p < .02$]. Overall, the overlearning group was more accurate than the underlearning group (76.3% vs. 69.5%). This main effect was qualified by an interaction between learning group and standard type [$F(1,18) = 5.24$, $p < .05$]. However, this interaction, surprisingly, ran counter to what might be expected. The overlearning group was slightly more accurate than the underlearning group for comparisons involving the CVC standards (69.4% vs. 65.5%) but was even more accurate than the underlearning group for the successive perceptual comparisons (83.3% vs. 73.4%). The reason for this latter result is not clear.[3] What is clear, however, is that the differential learning criterion had a minimal effect on performance with the remembered standards and can be attributed to a learning ceiling effect.

One final interaction involving learning group was obtained. This was a four-way effect involving standard type, standard length, and comparison pair [$F(2,36) = 4.48$, $p < .05$]. This complex interaction accounts, in part, for the main effect. The overlearning group was more accurate than the underlearning group for comparisons with the end pairs of the short perceptual standard (82.9% vs. 71.3%). Because there were no other interactions involving group, we combined the data over the two levels of this factor for the remaining analyses.

Psychometric-Function Analyses

Figure 1 provides the *z* score psychometric functions for the three levels of perceptual and remembered standards for the combined data in Experiment 1, and Table 1 provides the psychophysical-performance measures associated with these functions. Each plot was obtained by taking the first-order regression through the *z* scores that corresponded to the probability that the variable stimulus was judged longer than the standard.[4] Note that the *z* scores are plotted with respect to an *effect coding* (see Bock & Jones, 1968, p. 32) on the *x*-axis. That is, the four comparison stimuli, with values not equal to the standard, are converted to ordinal points (−2, −1, 1, 2) to permit direct comparison of the slopes (i.e., discriminative sensitivity) in the various conditions.

Discriminative sensitivity. The JND provides an index of discriminative sensitivity for each standard. The JNDs in Table 1 were obtained by taking one half of the difference between the upper and lower limen [JND = (UL − LL)/2] for each psychometric function, where the UL and LL are defined as the values of *x* corresponding to *z* scores of 0.675 and −0.675, respectively. The Weber fraction, or normalized sensitivity index, is also provided in Table 1, permitting an investigation of Weber's law for comparisons with the perceptual and remembered standards.

Figure 1 shows that the psychometric functions obtained with the perceptual standards display an increase in slope as standard length increases. This is evident as a decrease

Table 1
Point of Subjective Equality (PSE), Just Noticeable Difference (JND), and Weber Fraction (WF) for Comparison with the Perceptual (P) and Remembered (R) Standards in Experiment 1

Standard (mm)	PSE (mm)	JND (mm)	WF
10.54 (P)	10.57	.54	.051
(R)	11.11	.79	.076
100 (P)	99.68	4.13	.041
(R)	97.27	9.24	.092
200 (P)	199.58	6.48	.032
(R)	200.44	8.49	.042

of the Weber fraction with standard length, providing a violation of Weber's law. Specifically, fixed ratios require constancy in the slopes of the psychometric functions, and thus constant Weber fractions, for Weber's law to hold. Indeed, 17 of 20 subjects displayed a steeper psychometric function for the long standard than for the short standard.

In contrast to performance with the perceptual standards, discriminative sensitivity with the remembered standards is reduced. This is evident by shallower slopes of the psychometric functions and thus larger JNDs and Weber fractions for each standard value. Furthermore, comparisons involving the remembered standards show a similar violation of Weber's law. Here, 18 of 20 subjects displayed a steeper psychometric function for the long standard than for the short standard. Unlike with the perceptual standards, however, the Weber fraction does not show a consistent decrease with standard length; there is a substantial peak in the Weber fraction for the middle CVC standard, denoting a marked reduction in discriminability for comparisons made with that standard.

Individual-subject analyses. Psychometric functions were computed for each subject at each level of perceptual and remembered standard. The slopes[5] of these psychometric functions were subjected to a trend analysis with standard type (CVC, line) and the three standard lengths as within-subject factors. The linear component of standard length was highly reliable [$F(1,19) = 38.06, p < .0001$], confirming that the slopes of the psychometric functions increased as standard length increased and verifying the violation of Weber's law with the perceptual and remembered standards. The quadratic component was also reliable [$F(1,19) = 6.99, p < .02$], indicating that the psychometric-function slopes increase sharply for comparisons with the 200-mm standard. The main effect of standard type was highly reliable [$F(1,19) = 30.01, p < .0001$], confirming that the slopes of the psychometric functions were steeper with the perceptual than with the remembered standards. The linear component of the interaction between standard length and standard type was not reliable [$F(1,19) < 1.0$]. However, the quadratic component was reliable [$F(1,19) = 4.62, p < .05$], confirming the peak in the Weber fraction for comparisons with the middle CVC standard.

Finally, it is interesting to note the relationship between discriminative sensitivity with a perceptual and with a remembered standard. The Pearson product-moment correlation between the slopes of the individual-subject psychometric functions with the perceptual and remembered standards was .462 [$F(1,58) = 15.61, p < .0003$]. This suggests that the limits of resolution and distinctiveness of perceptual stimuli are closely related to the resolution and distinctiveness of remembered stimuli (cf. Shepard & Chipman's, 1970, "second order isomorphism").

Stimulus location precision. The PSEs for the combined data are provided in Table 1 and can also be seen in Figure 1 as the points at which the psychometric functions cut the x-axis. In contrast to the actual standard value, the PSE corresponds to the subjective standard value assumed by the observer. Thus, the difference between the actual standard and the PSE, commonly referred to as the constant error, provides an index of stimulus location precision.

Figure 2 provides frequency distributions of individual-subject PSEs for each level of perceptual and remembered standard. As expected, the PSEs for successive perceptual comparisons are virtually identical to the actual standard values. In contrast, the PSEs *representing* the magnitudes of the remembered standards show greater variability. In addition, consistent with the reduced discriminability evident for the middle CVC standard, 8 subjects provided PSEs outside the range of comparison stimuli for the middle standard as compared with 6 for the short standard and only 1 for the long standard.

Although the distributions of PSEs in Figure 2 provide an important view of the variability of long-term memories for percepts, it should be noted that deviations in the PSE can also arise as a consequence of simple response biases (see Engen, 1971). For example, a bias to view the memory standards as "longer" than the comparison stimuli would have the effect of shifting all the PSEs to the right. However, the bias interpretation loses considerable force when, as is evident in Figure 2, the distributions of PSEs are shifted to the right for the short standard, to the left for the middle standard, and are evenly distributed for the long standard. In addition, there were no consistent tendencies for individual subjects to either over- or underestimate the three standards.

One factor that provides a partial explanation for these PSE shifts is the classic notion of memory "pooling" or "central-tendency" (Hollingworth, 1910; Leuba, 1892). According to this view, items in long-term memory will exhibit a natural propensity to migrate, or drift, towards the geometric mean, or "adaptation-level" (Helson, 1947), of the subjective stimulus range. The geometric mean of the standards in the present study (10.54, 100, and 200 mm), following Helson (1947), is 59.3 mm. By this process, then, the short standard should exhibit a positive drift in the PSE and the mid- and long-range standards should exhibit a negative PSE drift. Although the PSE drifts for the short- and mid-range standards are consistent with the predictions of the central-tendency view, the PSEs for the long standard do not show any consistent drift tendencies.

Presentation-Order Analyses

Table 2 provides TOE indices for comparison with the three levels of perceptual and remembered standards. Following Jamieson and Petrusic (1975a, 1975c), the TOE index is defined as

$$\text{TOE} = \tfrac{1}{2}[(P(c)|L\text{-}S) - (P(c)|S\text{-}L)],$$

where L-S and S-L refer to the longer-stimulus-first and shorter-stimulus-first presentation orders, respectively.

With successive perceptual comparisons we see the classic negative TOE. That is, accuracy is higher when the

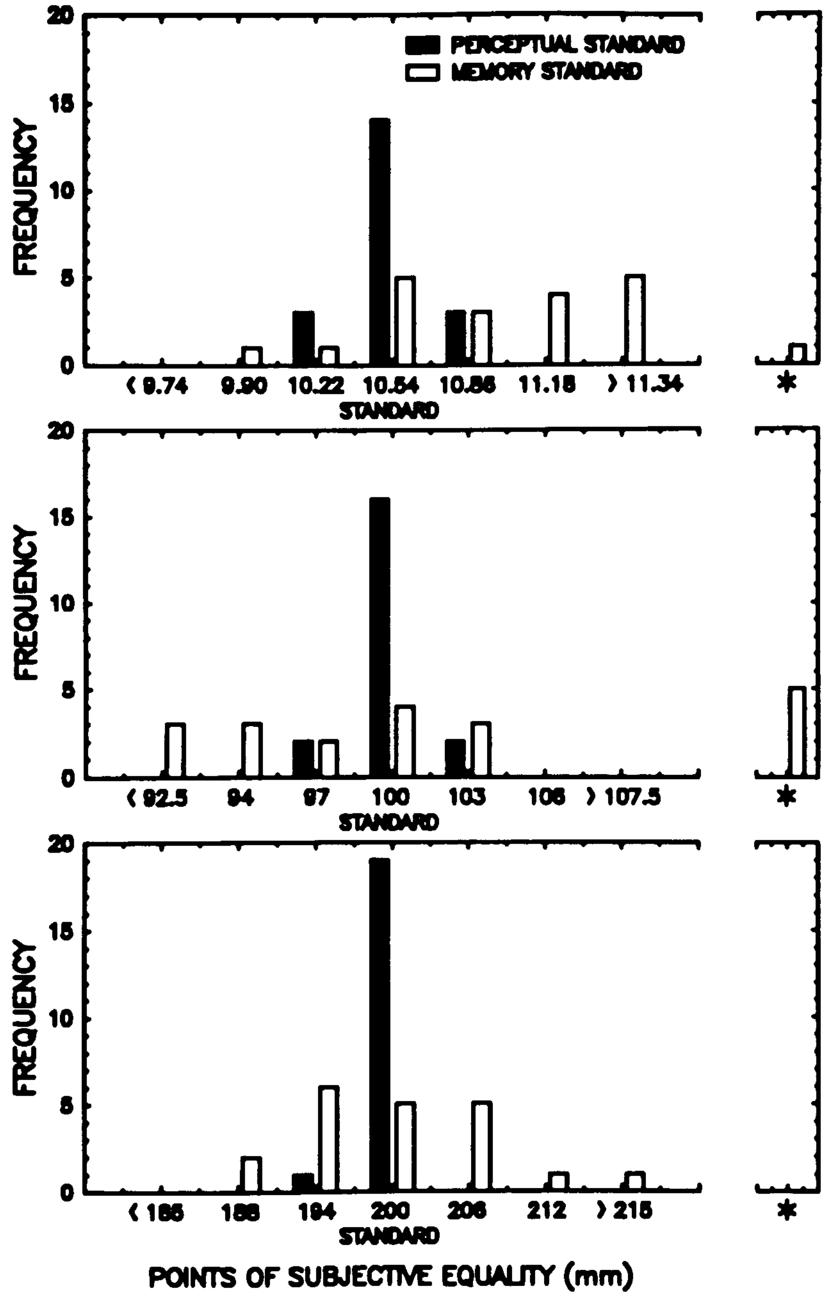

Figure 2. Frequency distributions of PSEs for the three levels of perceptual (filled bars) and remembered (open bars) standards in Experiment 1. Interval widths are .32, 3, and 6 mm for the 10.54-, 100-, and 200-mm standards, respectively. The symbol * denotes instances in which PSEs could not be calculated because the slope of the psychometric function approached zero or was negative.

longer line is presented second. The long perceptual standard does not show a TOE because comparisons made with that standard were very accurate and TOEs are known to decrease in magnitude as discriminability increases (e.g., Jamieson & Petrusic, 1975c).

In contrast, TOEs are clearly not evident for comparisons involving the remembered standards. The ANOVA reported in the first section of the Results revealed a reliable three-way interaction between standard type, comparison stimulus, and presentation order [$F(1,18) = 7.54$, $p < .02$], confirming that TOEs occurred only for comparisons involving the perceptual standards.

Comparisons involving identical pairs. Performance on trials involving comparisons of the standard with itself were consistent with the global effects of presentation order reported above. Specifically, for line standards, the second line was judged longer than the first on 55.2% of the trials [$\chi^2(1,N = 718) = 7.627$, $p < .01$], with the largest difference (56.1%) occurring with the short standard where the TOE was the largest and the smallest

Table 2
TOE Index (%) for the Three Levels of Perceptual and Remembered Standards in Experiment 1

	Perceptual	Remembered
Short	−7.7	−1.6
Medium	−4.5	−1.4
Long	+0.4	+2.1

difference (53.8%) occurring with the long standard where the TOE was negligible. In contrast, for standard-CVC or CVC-standard comparisons, the second stimulus was judged the longer on only 51.5% of the trials [$\chi^2(1,N = 717) = 0.615, p > .5$].

Response-Time Analyses

Figure 3 plots mean response times for comparisons with the three levels of perceptual and remembered standards as a function of presentation order (standard first vs. standard second) for the combined data of all subjects. As expected, response times are slower when comparison involves a remembered standard. In addition, the TOE evident with the percent-correct measure in the previous section is also evident with the response-time measure for comparisons involving the perceptual standards. Specifically, for variable stimuli less than the standard (i.e., 9.90, 10.22, 94, 97, 188, and 194 mm), response times are faster in the variable-standard presentation order because the standards are longer. On the other hand, response times are faster in the standard-variable presentation order for comparisons involving variable stimuli greater than the standard because the variable lines are longer.

For comparisons involving the remembered standards, there is clearly no comparable TOE effect on response time. However, there is evidence for a presentation-order effect of a different variety. Namely, response times are uniformly faster in the variable-standard presentation order (i.e., the line-CVC order). This result makes sense. If the CVC appears first, then the subjects must wait until the line is presented before they can initiate the comparison. On the other hand, if the line appears first, then the subjects immediately know which of the three CVCs *might* be presented (i.e., on two thirds of such trials, the subsequent stimulus will be another line) and they apparently initiate the appropriate retrieval operation during the ISI.

Although response times were faster in the line-CVC order, response accuracy was slightly poorer (65.5% vs. 69.3%). Rather than a strategic speed-accuracy tradeoff effect, the faster but less accurate responding in the line-CVC order is most likely due to early retrieval of the CVC standard and perceptual-memory decay of the line, respectively (further evidence for this view will be provided in Experiment 2).

Finally, collapsing the data over the two presentation orders reveals that response times for comparisons with both the perceptual and remembered standards were sensitive to the relative difficulty of the judgments. Specifically, in each case, response times peak at the PSE (see Table 1) and decrease as the difference between the PSE and variable stimuli increases.

A four-way ANOVA with mean response time as the dependent variable and standard type (line, CVC), the three standards, four comparison stimuli, and presentation order (S-L, L-S) as within-subject factors was used to evaluate the results. Response times were faster for comparisons with a perceptual standard than with a CVC standard [1,665 msec vs. 2,685 msec; $F(1,19) = 84.01, p < .0001$]. The main effect of standard length was reliable [$F(2,38) = 7.52, p < .002$]; comparisons involving the short standard were fastest for both the perceptual standards (short = 1,611 msec, medium = 1,691 msec, long = 1,692 msec) and the CVC standards (short = 2,536 msec, medium = 2,809 msec, long = 2,711 msec). The interaction between standard type and standard length was not

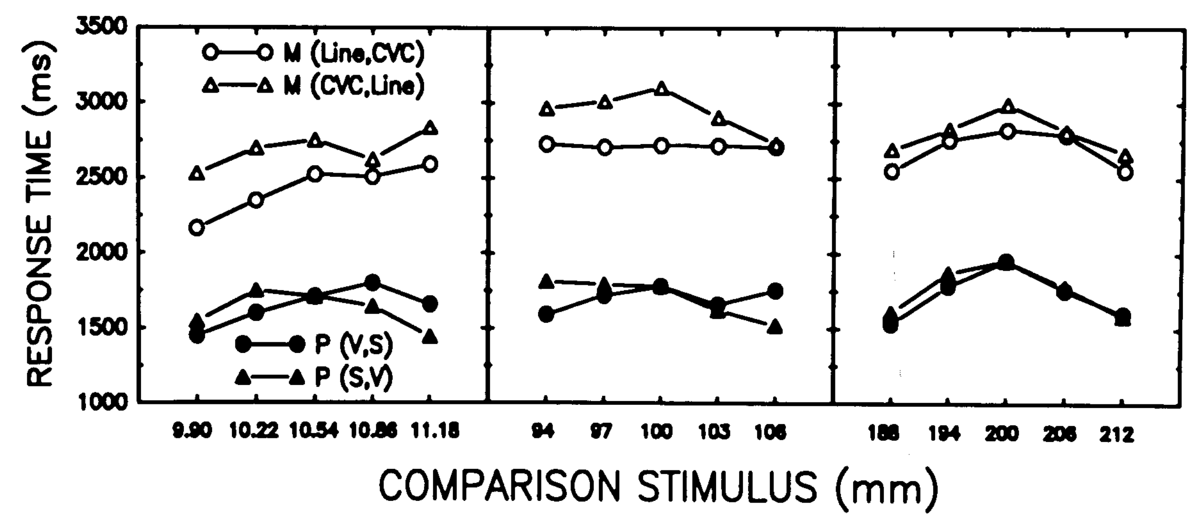

Figure 3. Mean response times for comparisons with the perceptual (P; filled points) and memorial (M; unfilled points) standards in Experiment 1. The variable-standard presentation order (i.e., V,S or Line,CVC) is denoted by circles, and the standard-variable presentation order (i.e., S,V or CVC,Line) is denoted by triangles.

reliable [$F(2,38) = 2.68, p > .08$]. The main effect of comparison stimulus was reliable [$F(3,57) = 4.39, p < .011$], which confirmed that response times depended on the values of the variable stimuli. The main effect of order was reliable [$F(1,19) = 4.61, p < .05$], as was the interaction between comparison pair and order [$F(3,57) = 3.86, p < .05$]. The latter effects were qualified by a three-way interaction between standard type, order of presentation, and comparison stimulus [$F(3,57) = 2.35, p < .04$], which confirms the response-time TOE for the successive perceptual comparisons and the difference between line-CVC and CVC-line comparisons for trials involving remembered standards.

Discussion

Three main findings were obtained in Experiment 1. First, it is possible to perform confusable relational judgments with remembered magnitudes and thus obtain the classic psychophysical-performance measures of JND, PSE, and Weber fractions. Although discriminative sensitivity is reduced when comparison involves a symbolic standard, comparisons with either a perceptual or a symbolic standard are sensitive to increases in stimulus differences with a common ratio (i.e., Weber's law is violated). Second, relative to performance with the short and long CVC standards, comparison with the middle CVC standard displayed a reduction in discriminative sensitivity (i.e., higher Weber fraction) and a reduction in location precision (i.e., more variable PSEs). Finally, comparisons involving symbolic standards are somehow distinct because they do not show a TOE.

It is interesting to note that Weber's law was violated in a similar way for comparisons with the perceptual and remembered standards (i.e., superior discriminative sensitivity with the long standard). However, further analyses revealed that the TOE was involved in the violation of Weber's law for the successive perceptual comparisons but not for comparisons involving the remembered standards. Figure 4 provides plots of psychometric functions for each level of perceptual and remembered standards as a function of presentation order (L-S vs. S-L). For comparisons with the perceptual standards (left column), the interesting result is that Weber's law actually holds remarkably well in the short-long presentation order. In the long-short presentation order, on the other hand, the slopes of the psychometric functions vary with standard length and the magnitude of the deviation is directly proportional to the magnitude of the TOE (see Table 2). Interestingly, Getty (1975) obtained a similar result in the context of successive comparisons of dura-

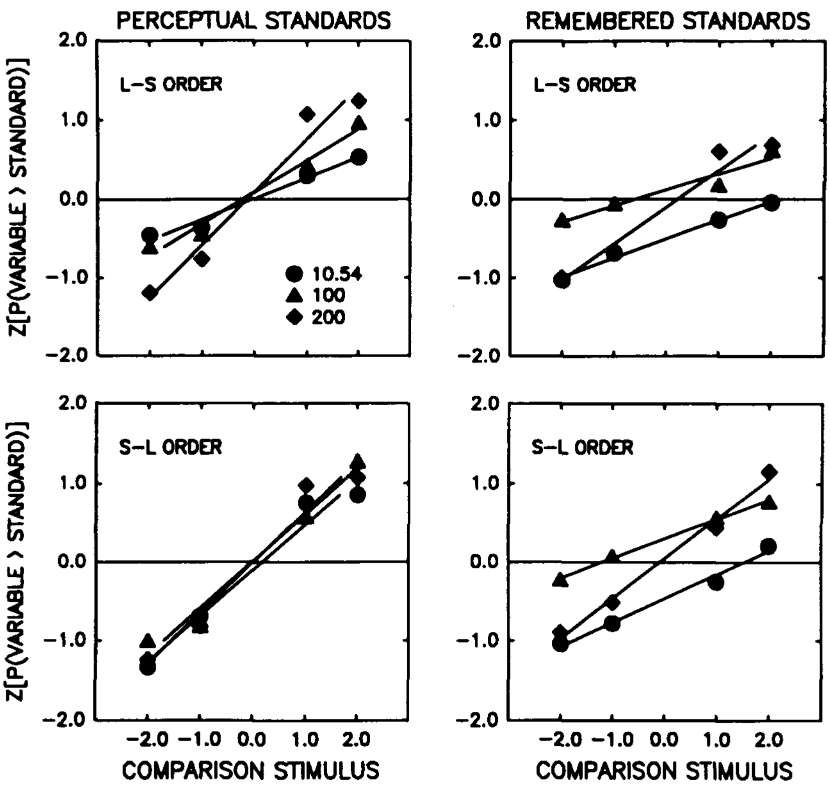

Figure 4. Effect-coded psychometric functions for the three levels of perceptual (left column) and remembered (right column) standards for the short-long (S-L) and long-short (L-S) presentation orders in Experiment 1.

tion discrimination. Getty found that Weber's law held for durations of the standard stimulus in the range from 200 to 2,000 msec. For durations beyond 2,000 msec, Getty obtained a dramatic increase in the Weber fraction and this violation was concomitant with the appearance of a large, negative TOE.

For comparisons involving the remembered standards, as previously noted, there is no difference in performance with respect to presentation order. In fact, the psychometric functions are virtually identical in the two orders, and the clear basis for the violation of Weber's law is the dramatically increased discriminative sensitivity with the long standard.

Before discussing the implications of these results as well as those stated above, we concede that after completing Experiment 1 we were concerned that the degree of location precision attained with the remembered standards could have been due, in part, to the intermixing of line-line comparisons, with comparisons involving CVCs. Specifically, one half of the trials involved line-line comparisons, and it might have been possible that repeated exposure to a limited range of variable stimuli allowed subjects to acquire some information about the value of the standard. However, bear in mind that subjects did not know that a member of a given line-line pair was a standard, let alone which one was the standard. In any case, the purpose of Experiment 2 was to replicate the major results obtained in Experiment 1 while removing the potential for any alternative strategies. In Experiment 2, we used the same task but had subjects participate in two sessions, the first involving only CVC-line and line-CVC comparisons and the second involving only line-line comparisons.

EXPERIMENT 2

Method

Subjects

Twelve Carleton University graduate students participated in two experimental sessions of approximately 1 h in duration. All subjects were naive with respect to the nature and aims of the experiment and were paid $15 for their participation.

Stimuli and Apparatus

The stimuli and apparatus were the same as in Experiment 1.

Procedure

Session 1 involved only CVC-line and line-CVC comparisons. The session was preceded by the overlearning phase described in Experiment 1. On Day 2, the subjects returned for a session of line-line comparisons. Both sessions consisted of four blocks of 60 randomized trials (3 standards × 5 comparison stimuli × 2 instructions × 2 presentation orders). All other aspects of the procedure were the same as in Experiment 1.

Results and Discussion

As in Experiment 1, trials on which response times exceeded 10 sec were excluded from the analyses. This accounted for 13 (0.2%) of the 5,760 trials obtained. The results are presented in three sections. The first examines the psychometric functions and associated indices of discrimination performance. The second investigates the effect of presentation order, and the third examines the response-time data for these comparisons.

Psychometric-Function Analyses

Figure 5 provides the psychometric functions for the combined data in Experiment 2, and Table 3 provides the

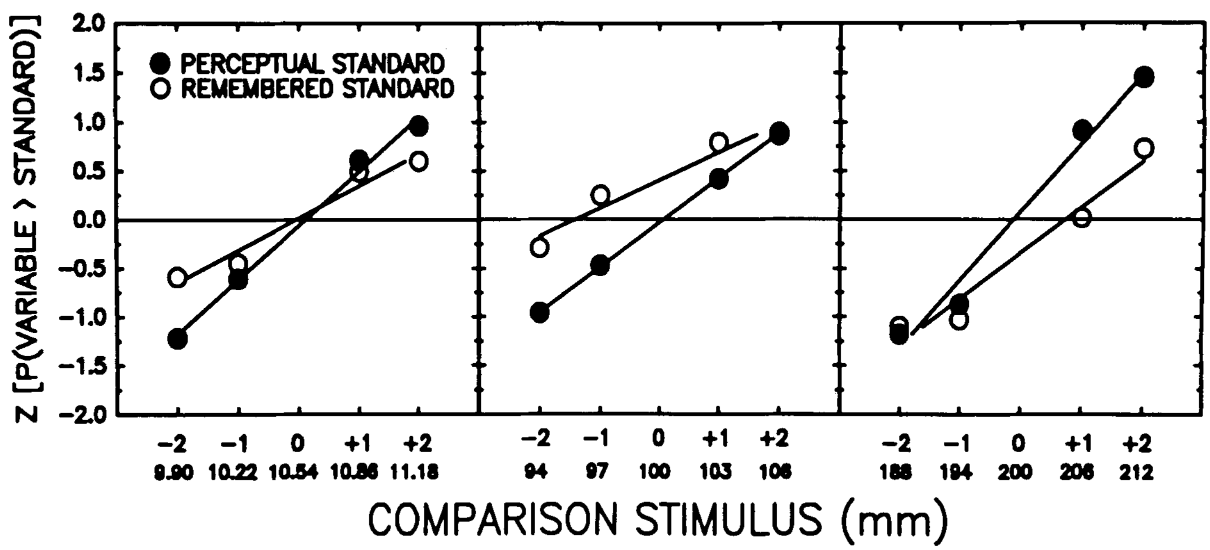

Figure 5. *Z* score psychometric functions for the three levels of perceptual (filled circles) and remembered (open circles) standards in Experiment 2. R^2 values for the 10.54-, 100-, and 200-mm standards are, respectively, .958, .939, and .953 for the remembered standards and .992, .999, and .980 for the perceptual standards.

Table 3
Point of Subjective Equality (PSE), Just Noticeable Difference (JND), and Weber Fraction (WF) for Comparison with the Perceptual (P) and Remembered (R) Standards in Experiment 2

Standard (mm)	PSE (mm)	JND (mm)	WF
10.54 (P)	10.58	.39	.037
(R)	10.53	.65	.062
100 (P)	99.80	4.11	.041
(R)	95.75	7.08	.071
200 (P)	199.34	5.75	.028
(R)	204.60	8.60	.043

psychophysical-performance measures associated with these functions.

Discriminative sensitivity. The main results reported in Experiment 1 were replicated: (1) Accuracy was higher for comparisons with the perceptual standards than with the CVC standards (80.0% vs. 70.1%), resulting in higher JNDs and Weber fractions for each memory standard; (2) the slopes of the psychometric functions increased with standard length for the perceptual and remembered standards, confirming the violation of Weber's law with each; specifically, for both types of standards, 9 of 12 subjects displayed a steeper psychometric function for the long standard than for the short standard; and (3) discriminative sensitivity was poorest for comparisons made with the middle CVC standard, evident in the inflated Weber fraction obtained for that standard.

Individual-subject analyses. A trend analysis paralleling the one in Experiment 1 was used to analyze the results. Again, the slopes of the individual-subject psychometric functions were used as the dependent variable, with standard type (line, CVC) and the three standard lengths as within-subject factors. The main effect of standard type was reliable [$F(1,11) = 30.37, p < .0001$], which confirmed that the Weber fractions were higher for comparisons with the remembered standards. A reliable effect of standard length [$F(2,22) = 4.97, p < .025$] confirmed the decrease in the Weber fraction with both the perceptual and remembered standards. Unlike in Experiment 1, however, the quadratic interaction between standard type and standard length was not reliable ($F < 1$) because the Weber fraction did not monotonically decrease for the perceptual standards. Rather, it actually increased slightly for the middle standard and then dropped off sharply for the long standard. As in Experiment 1, however, deviations in the Weber fraction for the successive perceptual comparisons were directly related to the TOE (see below).

Finally, as in Experiment 1, there was a strong correlation between the slopes of the individual-subject psychometric functions with the perceptual and remembered standards ($r = .608$, $F(1,34) = 19.91, p < .0001$].

Stimulus location precision. Figure 6 shows that the PSEs for the perceptual standards were virtually identical to the actual standards, while those representing the remembered standards were again highly variable. Furthermore, the reduced discriminative sensitivity for the middle CVC standard is also evident in the variability of the PSEs: 5 of 12 subjects provided PSEs outside the range of comparison stimuli with the mid-range standard as compared with 3 of 12 subjects with the short and long standards.

Also evident in Figure 6 is that only the mid-range standard displays a pattern of PSE drift that is consistent with the results reported in Experiment 1. Here, the PSEs are evenly distributed for the short standard, show a negative drift for the mid-range standard, but display a positive drift tendency for the long standard. Importantly, this latter result runs contrary to the predictions of the central-tendency view discussed in Experiment 1. Hence, when taken together, the results of the present studies provide very limited support for the notion of central tendency for remembered magnitudes. Instead, it is likely that these distributions simply reflect idiosyncratic differences (or biases) in estimation of the standards.

Presentation-Order Analyses

Table 4 provides the TOE indices for comparison with the three levels of perceptual and remembered standards.[6] The critical replication is that TOEs were not evident with the remembered standards while negative TOEs occurred for the successive perceptual comparisons. Note the especially large TOE for the middle standard. Here, accuracy was 92.4% in the S-L order but only 54.5% in the L-S order! Clearly, then, the nonmonotonicity in the Weber fraction for the mid-range perceptual standard discussed above was a direct result of this very large TOE.

In addition, note that the TOE for the short perceptual standard is small compared with the one obtained in Experiment 1. However, this is a bit misleading. Three of the 12 subjects displayed a large, *positive* TOE with the short standard. Without these subjects, the TOE index jumps to −7.1%. This is not to say that these subjects were devious. On the contrary, they actually display the classic result, known since Fechner (1860/1966), that TOEs are predominantly positive for small magnitudes and then reverse sign and become predominantly negative as magnitudes increase (cf. Petrusic & Baranski, 1989b, Experiment 2).

A four-way ANOVA with standard type, standard length, comparison stimulus, and presentation order (L-S, S-L) as within-subject factors was used to evaluate the effects of presentation order. The main effect of order was reliable [$F(1,11) = 40.12, p < .0001$], as were the interactions between order and comparison stimulus [$F(1,11) = 33.17, p < .0001$] and order and standard length [$F(2,22) = 5.53, p < .03$]. In addition, the three-way interaction between order, standard type, and standard length was reliable [$F(2,22) = 4.49, p < .05$], as was the three-way interaction between order, standard length, and comparison stimulus [$F(6,66) = 3.01, p < .05$]. The critical interaction between standard type, comparison stimulus, and order was highly reliable [$F(3,33) = 7.03$,

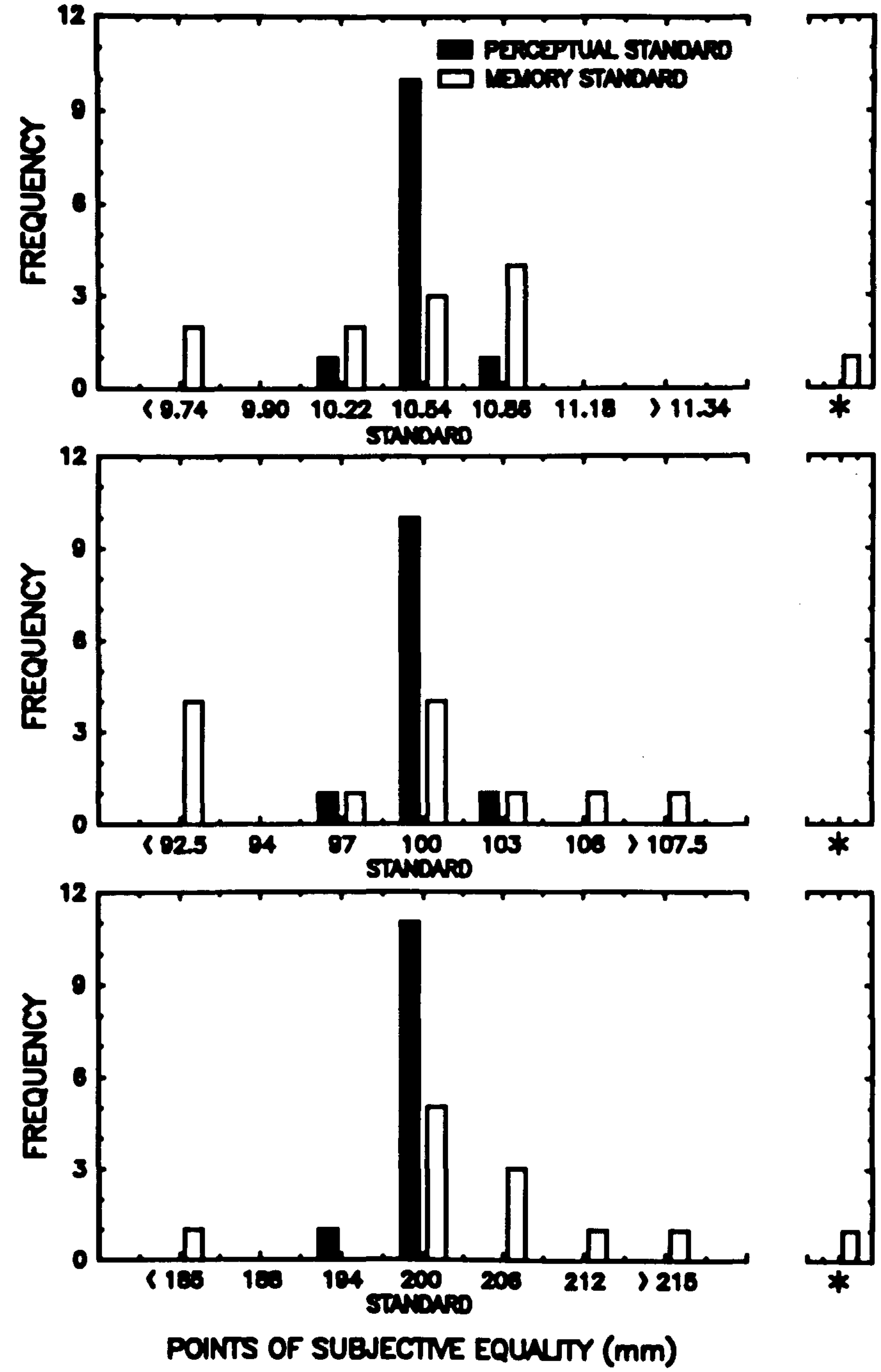

Figure 6. Frequency distributions of PSEs for the three levels of perceptual (filled bars) and remembered (open bars) standards in Experiment 2. Interval widths are .32, 3, and 6 mm for the 10.54-, 100-, and 200-mm standards, respectively. The symbol * denotes instances in which PSEs could not be calculated because the slope of the psychometric function approached zero or was negative.

$p < .002$], which confirmed the TOE for the successive perceptual comparisons and their absence with memory standards.

Comparisons involving identical pairs. As in Experiment 1, performance on trials involving comparison of the standard with itself was consistent with the global effects of presentation order. For successive line-standard comparisons, the second line was judged longer than the first on 65.9% of the trials [$\chi^2(1, N = 575) = 58.24$, $p < .01$], with the largest difference (75.9%) occurring with the medium-length standard where the TOE was the largest, and the smallest difference (58.3%) occurring with the short standard where the TOE was relatively small. On the other hand, for standard-CVC and CVC-standard

Table 4
TOE Index (%) for the Three Levels of Perceptual and Remembered Standards in Experiment 2

	Perceptual	Remembered
Short	−1.1 (−7.1*)	−0.2
Medium	−17.2	−3.5
Long	−9.3	+0.8

*TOE index excluding three subjects with positive TOEs.

comparisons, the second stimulus was judged the longer on only 52.2% of the trials [$\chi^2(1, N = 575) = 1.087$, $p > .1$].

TOEs and Weber's law. An analysis comparable to the one provided in Experiment 1 between presentation order and Weber's law is complicated by the fact that some subjects displayed positive TOEs with some standards while others showed negative TOEs with all standards. However, when we remove the subjects with positive TOEs, Weber's law again holds in the S-L presentation order as it did in Experiment 1. In the L-S order, however, the slope of the psychometric functions are shallowest for the middle standard because that is where the largest TOE was obtained. On the other hand, the psychometric functions for the remembered comparisons are very similar to those reported in Experiment 1; there is no effect of presentation order, and the clear violation of Weber's law occurs with the long standard in both presentation orders.

Evidently, there are very few conditions under which Weber's law can be studied for perceptual comparisons of linear extent. The simultaneous viewing condition does not provide a valid test because the task is akin to a visual acuity task (Ono, 1967). The successive comparison condition, on the other hand, is vulnerable to the effects of presentation order.

Response-Time Analyses

Figure 7 plots the mean response time for each level of perceptual and remembered standards for the two presentation orders for the combined data in Experiment 2. Comparisons involving the perceptual standards show a clear replication of the TOE effect on response time reported in Experiment 1. In addition, there was again no evidence of a comparable TOE effect for comparisons involving the remembered standards. However, the same CVC-line versus line-CVC order effect reported in Experiment 1 is again evident; this time it is even greater in magnitude. Here, response times in the line-CVC presentation order are about the same as for the successive *perceptual* comparisons and the difference between the CVC-line and line-CVC orders is approximately 800 msec (2,149 msec vs. 1,334 msec)! The reason for this enhanced effect is that, in this experiment, if a line is presented first, then the subject knows that a CVC will follow and also knows which CVC will follow. As in Experiment 1, comparisons were slightly more accurate in the CVC-line order than in the line-CVC order (71.6% vs. 68.6%). Clearly, this effect is not due to a speed-accuracy tradeoff, because if it were, then accuracy would have been drastically reduced in the present experiment. Rather, it seems that the 3%-4% reduction in discriminative accuracy in the line-CVC order, in both experiments, is due to perceptual-memory decay of the line. Finally, as in Experiment 1, response times peak at the PSE for each level of perceptual and remembered standards (see Table 3) and decrease as the difference between the PSE and comparison stimuli increases.

A four-way ANOVA paralleling the one reported in Experiment 1 was used to evaluate the results. Response times were faster for comparisons involving the percep-

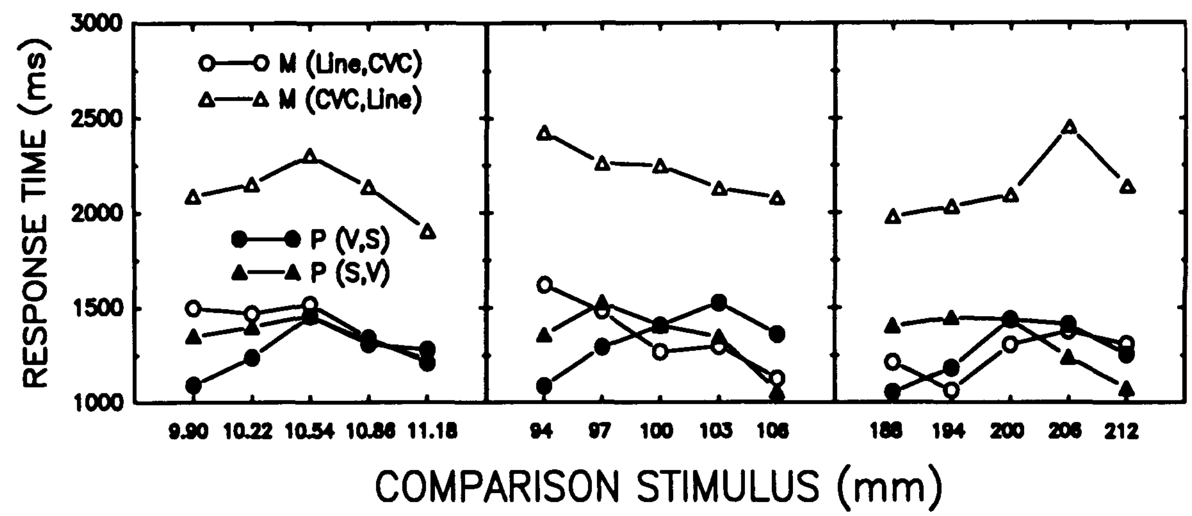

Figure 7. Mean response times for comparisons with the perceptual (P; filled points) and memorial (M; unfilled points) standards in Experiment 2. The variable-standard presentation order (i.e., V,S or Line,CVC) is denoted by circles, and the standard-variable presentation order (i.e., S,V or CVC,Line) is denoted by triangles.

tual standards [$F(1,11) = 13.03$, $p < .005$], depended on the value of the comparison stimulus [$F(3,33) = 6.18$, $p < .005$], but did not depend on the length of the standard ($p > .1$). Overall mean response times for the short, medium, and long standards were 1,277, 1,319, and 1,259 msec for the perceptual standards and 1,725, 1,803, and 1,696 msec for the CVC standards. The main effect of order was reliable [$F(1,11) = 5.43$, $p < .05$], as were the interactions between standard type and order [$F(1,11) = 34.40$, $p < .0001$] and comparison pair and order [$F(3,33) = 79.38$, $p < .0001$]. The latter effects were qualified by a highly reliable interaction between type of standard, comparison stimulus, and presentation order [$F(3,33) = 66.53$, $p < .0001$], which confirms the TOE for line-line comparisons but the dependence of line-CVC versus CVC-line order for comparisons involving the remembered standards.

GENERAL DISCUSSION

The present studies demonstrated that people can perform confusable relational judgments with symbolic representations of perceptual magnitudes. Although it is tempting to assume that this necessarily implies a detailed, analogue memory representation, it is important to first consider an alternative view. For example, it is possible to develop a plausible account of comparisons between a line and CVC in the present task in the context of the *semantic-coding* theory of Banks (1977; Banks et al., 1982; see also Shoben, Cech, Schwanenflugel, & Sailor, 1989, for a similar view). The semantic-coding theory postulates that the presentation of a CVC generates a discrete, categorical semantic code (e.g., "long"). We can refer to codes generated at this level as *superordinate* semantic codes. Comparison between a CVC and a line requires the generation of a categorical semantic code for the *line* (e.g., "long") to permit comparison in a common format (see Banks et al., 1982). Since this level of semantic coding will be insufficient to perform the comparison, a secondary, *subordinate* categorical code must be generated (e.g., "short" or "long"). To obtain psychometric functions with non-zero slope, then, the subject requires only very general hypotheses about the lengths of the standards and a process of hierarchical, probabilistic, binary categorization of the comparison *lines* within each subrange.

For example, as the subject becomes aware of the variable stimuli within each subrange, one of these variable stimuli is arbitrarily selected as the "standard." Each variable stimulus is then categorized as "longer" or "shorter," relative to the corresponding "standard." Of course, according to the semantic-coding theory, the categorization process is probabilistic depending on how well the variable stimuli can be ordered.

According to this extended semantic-coding view, the observed response-time difference between the line-CVC and CVC-line presentation orders arises not as a consequence of a retrieval time advantage in the line-CVC order but because the categorization process for the line can be initiated sooner in that order. Finally, this semantic-coding approach can also account for the observed violation of Weber's law for comparisons with CVC standards. Specifically, since the absolute stimulus difference between comparison lines increases as the length of the standard increases (i.e., .32, 3, and 6 mm), it can be assumed that probabilistic identification of these comparison lines, and thus subsequent categorical semantic coding of these lines, will be more accurate as the length of the standard increases. Herein lies the downfall of this model, however, because such a view must predict that discriminability will be higher for the middle standard than for the short standard since the spacing of the comparison lines is increased. On the other hand, the present experiments showed that there is a marked reduction in discriminability for comparisons with the middle CVC standard. Hence, even under very general assumptions, a discrete semantic-coding approach is insufficient for the present data.

The semantic-coding theory also proves insufficient as a general model of the comparison process because it was not designed to be a model for perceptual-comparison data. Given the many similar properties of the comparison process with perceptual and remembered/symbolic magnitudes (e.g., distance effects, semantic congruity effects; see Petrusic & Baranski, 1989a, 1989b, 1992), it is more parsimonious to assume a more general model of the comparison process that subsumes an analogue scale representation and accounts for the full range of perceptual and symbolic comparison data (e.g., Birnbaum & Jou, 1990; Holyoak, 1978; Link, 1990; Petrusic, in press).

Positional Effects in Symbolic Comparisons

In both experiments, the mid-range CVC standard displayed reduced discriminative sensitivity (i.e., inflated Weber fractions) and reduced location precision (i.e., variable PSEs). We believe that this result provides a new and revealing characterization of the *serial-position effect* in symbolic comparisons.

The term "serial-position effect" has been used in several distinct areas of experimental psychology; in each case denoting a specific performance deficit observed for central items in a set of ordered stimuli. Perhaps the best-known use of the term is in the context of the classic serial-list-learning tasks (see Johnson, 1991, for a recent review of the properties of serial-list learning). However, the term "serial-position effect" has also been used to denote poorer performance for central items in psychophysical tasks requiring judgments of positional order (Holyoak & Patterson, 1981) and stimulus identification (see Vickers, 1979, for a review, and see Marley & Cook, 1984, who refer to it as an *end anchor effect*).

In the context of symbolic comparison, the term "serial-position effect" has been used to characterize an inverted "U" or "bowed" response-time relation, suggesting that pairs near the ends of the continuum are compared faster than are those toward the middle of the range (see Banks, 1977; Banks, White, & Mermelstein, 1980; Holyoak & Patterson, 1981; Moyer & Dumais, 1978; and Shoben et al., 1989, for reviews of serial-position effects in symbolic comparisons, and see Birnbaum & Jou, 1990, who refer to it as an *end effect*). Moreover, the serial-position effect has been observed only for the response-time measure because comparisons are made between remembered items and thus must necessarily be supraliminal. Because comparisons are made between remembered items, the serial-position effect in symbolic comparison tasks is often unavoidably confounded by distance effects. For example, in the simple case of three remembered magnitudes, the end magnitudes are compared half of the time with an item one ordinal step away and half of the time with the item two ordinal steps away. On the other hand, the middle item is always compared with an item one ordinal step away. In contrast, each symbolic comparison in the present task involves a highly distinctive remembered magnitude, and because comparisons between standards are not required, there are no comparable distance effects operating. Rather, the serial-position effect obtained with the Weber fraction in the present studies implies a reduction in *discriminability* for the middle CVC standard, unconfounded by distance effects.

One view of the serial-position effect in symbolic comparison that is consistent with the present findings is the *positional discriminability* model of Holyoak and Patterson (1981). According to this model, the subjective magnitude of a stimulus is represented by a distribution of analogue position codes along a continuum with finite mean and variance. The serial-position effect is predicted because the distribution of positional codes will overlap more for mid-range items than for end items; that is, middle items will be less distinctive (Murdock, 1960). A further assumption made by Holyoak and Patterson, one that is particularly relevant to the present data, is that mid-range remembered magnitudes possess greater variance in the distribution of positional codes (p. 1298). That is, the *quality* of the representation is reduced, and there is greater positional uncertainty about the true remembered magnitude. According to this view, then, the positional effect obtained in the present studies (i.e., inflated Weber fractions and more variable PSEs for the middle CVC-standard) is a consequence of the mid-range comparison lines being compared with a "noisier" memory representation. Indeed, the present findings provide direct support for this assumption since, given the very wide spacing between the memory standards in the present task, it is unlikely that errors occurred because of retrieval difficulties or comparisons with an inappropriate standard.

The basis for the greater positional variance for mid-range remembered magnitudes can be understood in the context of Johnson's (1991) *distinctiveness* model of serial learning. According to this view, the relative distinctiveness of each item in a memory set is determined by two factors. The first, following Murdock (1960), is a distinctiveness measure, proportional to the distance between a particular stimulus and the geometric mean, or adaptation level (Helson, 1947, 1964), of all stimuli in the memory set. The second factor depends on the strength and number of forward and backward associations between each item in the set, where the strength between any two items is inversely proportional to the distance between the items and the amount of learning.

An important prediction of this view is that factors that influence performance in serial-list-learning tasks will also influence discriminative performance with remembered standards in the present paradigm. Such factors include the number of standards, their relative spacing, and the number of learning trials provided. As previously mentioned, the ineffectiveness of the differential-learning manipulation employed in Experiment 1 was most likely a consequence of a ceiling effect. Specifically, including trials to criterion, the "underlearning" group had 800% overlearning with a small and widely spaced stimulus set. Indeed, the effects of learning might become especially evident when more standards are introduced or when the spacing between standards is not as wide as was the case in the present studies.

Finally, it is important to note that the above predictions are consistent with the assumption of an analogue memory representation. Conversely, it is not at all clear as to why factors such as the spacing of the memory standards, the number of standards, and the amount of learning should influence discriminative performance in the context of an ordinal, semantic-coding approach to comparisons between a line and a CVC in the present task. Hence, future studies with the present paradigm will most likely provide informative evaluations of these two general views of symbolic comparison.

The Absence of TOEs with Symbolic Magnitudes

As mentioned at the outset, several accounts of the TOE view the effect in terms of a perceptual-memory phenomenon (e.g., Hellström, 1985; Jamieson & Petrusic, 1975a, 1975c) and, as such, the present findings provide support for these views. On the other hand, the failure to obtain TOEs with symbolic magnitudes is inconsistent with models that conceptualize the effect in terms of central decisional-criterion biases (Allan, 1979; Luce & Galanter, 1963) or peripheral-response biases (Engen, 1971); otherwise TOEs would have been evident for all types of comparisons in the present studies. More generally, the failure to obtain TOEs with symbolic magnitudes suggests that the representations of perceived and remembered stimuli differ at some point in the information-processing stream.

To compare a memory with a percept requires an interface at some point (i.e., a common or highly similar representational format). Moreover, this interface must be, at the latest, at the decision/comparison stage because, as mentioned above, comparisons involving a perceived and a remembered stimulus exhibit the same properties of the comparison process as when two symbolic or two perceptual magnitudes are compared. In addition, we also know that TOEs are not necessarily a property of the comparison process because they have also been observed in tasks that do not require comparison or discrimination. For example, TOEs are clearly evident in tasks involving the *reproduction* of stimulus duration intervals (e.g., Eisler, 1975; Jamieson & Petrusic, 1975b; Petrusic, 1984; Schab & Crowder, 1988).

Taken together, these results implicate a *predecisional* locus for the TOE. Specifically, in line with Hellström's (1985) differential-sensation weighting theory, it seems that the TOE is a consequence of the pairing of successive perceptual stimuli, which, depending on stimulus magnitude and ISI, leads to the *subsequent* comparison of perceptual-memory representations that have either been diminished (assimilation) or accentuated (contrast). If TOEs do not occur with symbolic representations, then these representations must either exist, predecisionally, in a format that is not subject to sensation weighting or else enter into the information-processing stream, in a similar format, at some point after sensation weighting has taken place. In either case, it seems that a remembered magnitude does not exhibit the full range of functionally equivalent perceptual-like properties.

It is important to note that this view follows from the assumption of successive stimulus presentation in the present task. Conversely, it could be argued that TOEs were not evident with CVC standards because those comparisons were simply not successive. For example, in the CVC-line order, subjects might retrieve the appropriate magnitude and then compare it with the physically present line. Similarly, the faster response times in the line-CVC order suggest that retrieval of the standard is taking place while the line is still available.

However, although response times were relatively fast in the line-CVC order, they were not faster than the line-line comparisons. The line-CVC comparisons would have to be considerably faster than the line-line comparisons before we could assert that the latter were completed before the removal of the line. Since line-CVC comparisons took much longer than line-line comparisons in Experiment 1 and were not faster than the line-line comparisons in Experiment 2, despite a 500-msec response-time advantage during the ISI, the possibility is unlikely. In addition, we also know that comparisons in the line-CVC order were less accurate than in the CVC-line order (3%-4%), which indicates that the line is indeed undergoing perceptual-memory decay before the completion of the comparison. Hence, although it is possible that comparisons in the CVC-line order were simultaneous, there is evidence for successive comparison in the line-CVC order but no evidence of a TOE in that order.

In conclusion, if the symbolic comparisons in the present experiments were simultaneous, then Woodworth (1938) may still be correct in his conjecture after all. On the other hand, if the comparisons were indeed successive, then the effect of stimulus presentation order—one of the oldest problems in psychology—provides an important locus and criterion by which perception and memory can be distinguished.

REFERENCES

ALGOM, D., WOLF, Y., & BERGMAN, B. (1985). Integration of stimulus dimensions in perception and memory: Composition rules and psychophysical relations. *Journal of Experimental Psychology: General*, **114**, 451-471.

ALLAN, L. G. (1979). The perception of time. *Perception & Psychophysics*, **26**, 340-354.

ALLAN, L. G., & KRISTOFFERSON, A. B. (1974). Psychophysical theories of duration discrimination. *Perception & Psychophysics*, **16**, 26-34.

ANDERSON, N. H. (1970). Functional measurement and psychophysical judgment. *Psychological Review*, **77**, 153-170.

BANKS, W. P. (1977). Encoding and processing of symbolic information in comparative judgment. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 11, pp. 101-159). New York: Academic Press.

BANKS, W. P., MERMELSTEIN, R., & YU, H. K. (1982). Discriminations among perceptual and symbolic stimuli. *Memory & Cognition*, **10**, 265-278.

BANKS, W. P., WHITE, H., & MERMELSTEIN, R. (1980). Position effects in comparative judgments of serial order: List structure vs. differential strength. *Memory & Cognition*, **8**, 623-630.

BERKSON, J. (1953). A statistically precise and relatively simple method of estimating the bio-assay with quantal response, based on the logistic function. *Journal of the American Statistical Association*, **48**, 565-599.

BIRNBAUM, M. H., & JOU, J. (1990). A theory of comparative response times and "difference" judgments. *Cognitive Psychology*, **22**, 184-210.

BJORKMAN, M., LUNDBERG, I., & TARNBLOM, S. (1960). On the relationship between percept and memory: A psychophysical approach. *Scandinavian Journal of Psychology*, **1**, 136-144.

BOCK, R. D., & JONES, L. V. (1968). *The measurement and prediction of judgment and choice*. San Francisco: Holden-Day.

BRADLEY, D. R., & VIDO, D. (1985). Psychophysical functions for perceived and remembered distance. *Perception*, **13**, 315-320.

CATTELL, J. M. (1902). The time of perception as a measure of differences in intensity. *Philosophische Studien*, **19**, 63-68.

CHEW, E. I., & RICHARDSON, J. T. E. (1980). The relationship between perceptual and memorial psychophysics. *Bulletin of the Psychonomic Society*, **16**, 25-26.

EISLER, H. (1975). Subjective duration and psychophysics. *Psychological Review*, **22**, 429-450.

ENGEN, T. (1971). Psychophysics: 1. Discrimination and detection. In J. W. Kling & L. A. Riggs (Eds.), *Woodworth and Schlosberg's experimental psychology* (3rd ed., pp. 11-46). New York: Holt, Rinehart & Winston.

FARAH, M. J. (1985). Psychophysical evidence for a shared representational medium for mental images and percepts. *Journal of Experimental Psychology: General*, **114**, 91-103.

FECHNER, G. T. (1966). *Elements of psychophysics* (H. E. Adler, Trans.). New York: Harper & Row. (Original work published 1860)

FINKE, R. A. (1980). Levels of equivalence in imagery and perception. *Psychological Review*, **87**, 113-132.

FINKE, R. A. (1985). Theories relating mental imagery to perception. *Psychological Bulletin*, **98**, 236-259.

FINKE, R. A., & SHEPARD, R. N. (1985). Visual functions of mental imagery. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and performance: Vol. II: Cognitive processes and performance* (Chap. 37, pp. 1-55). New York: Wiley.

GETTY, D. J. (1975). Discrimination of short temporal intervals: A comparison of two models. *Perception & Psychophysics*, **18**, 1-8.

HEGELMAIER, F. (1852). Über des Gedächtnis für Linearanschauungen. *Archiv für Physiologisch Heilkunde*, **11**, 844-853.

HELLSTRÖM, A. (1977). Time errors are perceptual: An experimental

investigation of duration and a quantitative successive comparison model. *Psychological Research*, **39**, 345-388.
HELLSTRÖM, A. (1985). The time-order error and its relatives: Mirrors of cognitive processes in comparing. *Psychological Bulletin*, **97**, 35-61.
HELSON, H. (1947). Adaptation-level as a frame of reference for prediction of psychophysical data. *American Journal of Psychology*, **60**, 1-29.
HELSON, H. (1964). *Adaptation-level theory*. New York: Harper & Row.
HENDERSON, J. M., & WELL, A. D. (1985). Symbolic comparisons with and without perceptual referents: Is interval information used? *Memory & Cognition*, **13**, 176-182.
HENMON, V. A. C. (1906). The time of perception as a measure of differences in sensation. *Archives of Philosophy, Psychology & Scientific Methods*, **8**, 1-75.
HOLLINGWORTH, H. L. (1910). The central tendency of judgment. *Journal of Philosophy, Psychology & Scientific Methods*, **7**, 461-469.
HOLYOAK, K. J. (1978). Comparative judgments with numerical reference points. *Cognitive Psychology*, **10**, 203-243.
HOLYOAK, K. J., & PATTERSON, K. K. (1981). A positional discriminability model of linear-order judgments. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 1283-1302.
HULSE, S. H., DEESE, J., & EGETH, H. (1975). *The psychology of learning* (4th ed.). New York: McGraw-Hill.
JAMIESON, D. G., & PETRUSIC, W. M. (1975a). The dependence of time order error direction on stimulus range. *Canadian Journal of Psychology*, **29**, 175-182.
JAMIESON, D. G., & PETRUSIC, W. M. (1975b). Pairing effects and time-order errors in duration discrimination. *Perception & Psychophysics*, **18**, 107-113.
JAMIESON, D. G., & PETRUSIC, W. M. (1975c). Presentation order effects in duration discrimination. *Perception & Psychophysics*, **17**, 197-202.
JOHNSON, G. J. (1991). A distinctiveness model of serial learning. *Psychological Review*, **98**, 204-217.
KERST, S. M., & HOWARD, J. H., JR. (1978). Memory psychophysics for visual area and length. *Memory & Cognition*, **6**, 327-335.
KINCHLA, R. A., & SMYZER, F. (1967). A diffusion model of perceptual memory. *Perception & Psychophysics*, **2**, 219-229.
KOSSLYN, S. M. (1980). *Image and mind*. Cambridge: Harvard University Press.
LAMING, D., & SCHEIWILLER, P. (1985). Retention in perceptual memory: A review of models and data. *Perception & Psychophysics*, **37**, 189-197.
LEUBA, J. H. (1892). A new instrument for Weber's law: With indications of a law of sense memory. *Journal of Psychology*, **5**, 370-384.
LINK, S. W. (1990). Modeling imageless thought: The relative judgment theory of numerical comparisons. *Journal of Mathematical Psychology*, **34**, 2-41.
LUCE, R. D., & GALANTER, E. (1963). Discrimination. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 191-243). New York: Wiley.
MARLEY, A. A. J., & COOK, V. T. (1984). A fixed rehearsal capacity interpretation of limits on absolute identification performance. *British Journal of Mathematical & Statistical Psychology*, **37**, 136-151.
MILLER, G. A. (1947). Sensitivity to changes in the intensity of white noise and its relation to masking and loudness. *Journal of the Acoustical Society of America*, **19**, 609-619.
MOYER, R. S. (1973). Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception & Psychophysics*, **13**, 180-184.
MOYER, R. S., & BAYER, R. H. (1976). Mental comparisons and the symbolic distance effect. *Cognitive Psychology*, **8**, 228-246.
MOYER, R. S., BRADLEY, D. R., SORENSEN, M. H., WHITING, J. D., & MANSFIELD, D. P. (1978). Psychophysical functions for perceived and remembered size. *Science*, **200**, 330-332.
MOYER, R. S., & DUMAIS, S. T. (1978). Mental comparison. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 12, pp. 117-155). New York: Academic Press.
MOYER, R. S., & LANDAUER, T. K. (1967). Time required for judgments of numerical inequality. *Nature*, **215**, 1519-1520.
MÜNSTERBERG, H. (1894). Studies from the Harvard psychological laboratory: A psychometric investigation of the psychophysic law. *Psychological Review*, **1**, 45-51.
MURDOCK, B. B. (1960). The distinctiveness of stimuli. *Psychological Review*, **67**, 16-31.
ONO, H. (1967). Difference threshold for stimulus length under simultaneous and nonsimultaneous viewing conditions. *Perception & Psychophysics*, **2**, 201-207.
PETRUSIC, W. M. (1984). Explicit counting and time-order errors in duration discrimination. *Timing & Time Perception*, **423**, 630-633.
PETRUSIC, W. M. (in press). Semantic congruity effects and theories of the comparison process. *Journal of Experimental Psychology: Human Perception & Performance.*
PETRUSIC, W. M., & BARANSKI, J. V. (1989a). Context, context shifts, and semantic congruity effects in comparative judgments. In D. Vickers & P. Smith (Eds.), *Human information processing: Measures, mechanisms, and models* (pp. 231-251). Amsterdam: Elsevier/North-Holland.
PETRUSIC, W. M., & BARANSKI, J. V. (1989b). Semantic congruity effects in perceptual comparisons. *Perception & Psychophysics*, **45**, 439-452.
PETRUSIC, W. M., & BARANSKI, J. V. (1992). *On the properties of long term memories for percepts: Binary and quaternary relational judgments with remembered magnitudes*. Manuscript submitted for publication.
PHILLIPS, L. (1958). Mediated verbal similarity as a determinant of the generalization of a conditioned GSR. *Journal of Experimental Psychology*, **55**, 56-62.
SCHAB, F. R., & CROWDER, R. G. (1988). The role of succession in temporal cognition: Is the time-order error a recency effect of memory? *Perception & Psychophysics*, **44**, 233-242.
SHEPARD, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, **91**, 417-447.
SHEPARD, R. N., & CHIPMAN, S. (1970). Second order isomorphisms of internal representations: Shapes of states. *Cognitive Psychology*, **1**, 1-17.
SHOBEN, E. J., CECH, C. G., SCHWANENFLUGEL, P. J., & SAILOR, K. M. (1989). Serial position effects in comparative judgments. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 273-286.
STEVENS, S. S. (1957). On the psychophysical law. *Psychological Review*, **64**, 153-181.
VICKERS, D. (1979). *Decision processes in visual perception*. New York: Academic Press.
WOLF, Y., & ALGOM, D. (1987). Perceptual and memorial constructs in children's judgments of quantity: A law of across-representation invariance. *Journal of Experimental Psychology: General*, **116**, 381-397.
WOODWORTH, R. S. (1938). *Experimental psychology*. New York: Holt.

NOTES

1. The confidence data were obtained for a different purpose and are not reported here but are available through correspondence.

2. To obtain more reliable estimates of performance for the ANOVA, the four comparison pairs, with values not equal to the standard, were collapsed into two. Thus, the two pairs with comparison values less than the standard and the two pairs with comparison values greater than the standard were combined. In addition, p values for all analyses reported in this paper are with the Greenhouse-Geisser epsilon adjusted degrees of freedom. However, degrees of freedom reported in the text are not adjusted.

3. One possibility is that the additional learning trials actually increased the sensitivity of the comparison process for the perceptual stimuli. Indeed, there is evidence for an *acquired distinctiveness* with repeated exposure to perceptual stimuli, and the interested reader is referred to Hulse, Deese, and Egeth (1975, Chap. 8) for a review of this literature.

4. We omit points on the psychometric function for comparisons of the standard with itself because there is no "variable" stimulus in the successive line condition (i.e., unless $p(c) = .5$, the data points are am-

biguous). Rather, the PSE is derived directly from the psychometric function (see Bock & Jones, 1968, pp. 30-33). A summary of performance for comparisons involving identical pairs is provided later in the Results section under the heading "Presentation-Order Analyses."

5. In 6 of 120 instances, subjects displayed near-zero or slightly negative slopes to their psychometric functions (i.e., no discriminative sensitivity). All 6 were for comparisons with remembered standards, and 5 of the 6 instances occurred with the middle standard, again illustrating the reduced discriminability for that standard. Importantly, when the slope of the psychometric function is near zero or negative, PSEs, JNDs, and Weber fractions cannot be computed. For this reason, the statistical test for discriminative sensitivity was conducted directly on the slopes of the individual-subject psychometric functions. This analysis is appropriate because the JND and Weber fraction depend directly on the slope of the psychometric function.

Let $z = a+bx_1$ denote the z score psychometric function, where a is the intercept, b is the slope, and x_1 is the effect-coded variable stimulus $(-2, -1, 0, 1, 2)$. Let $x_2 = S(1+cx_1)$ be a linear transformation equation relating the effect-coded equation to one representing the physical stimulus values, where S is the standard and c is a conversion constant ($c = .03$ in the present case). Since $x_1 = (x_2-S)/Sc$, it can be shown that the upper limen (UL) $= Sc/b(.675-a+b/c)$ and that the lower limen (LL) $= Sc/b(-.675-a+b/c)$. Since JND = (UL−LL)/2, it follows that JND $= (.675)(S)(c)/b$ and the Weber fraction $= (.675)(c)/b$. Thus, the Weber fraction is inversely related to the slope, and constancy of slopes is required for the strong form of Weber's law to hold.

Finally, when calculating individual-subject z score psychometric functions, we corrected probabilities equal to 0 and 1.0 according to Berkson (1953), whose method is recommended by Bock and Jones (1968, p. 21).

6. Both here and in the first experiment, TOE direction was consistent over the two instructions (longer, shorter). This replicates previous research in the context of duration discrimination (Jamieson & Petrusic, 1975a, 1975c; Petrusic, 1984) and confirms that TOEs are not a result of peripheral response biases; otherwise the TOE would reverse with instructions.

(Manuscript received November 13, 1989;
revision accepted for publication September 23, 1991.)