Differential effects of stimulus context on perceived length: Implications for the horizontal–vertical illusion

LAURA ARMSTRONG and LAWRENCE E. MARKS
John B. Pierce Laboratory, New Haven, Connecticut
and Yale University, New Haven, Connecticut

Six experiments examined orientation-specific effects of stimulus context on the visual perception of horizontal and vertical lengths: Using a paired-comparison method, Experiments 1–5 showed that the probability of judging a given vertical line to be longer than a given horizontal line was relatively great when the stimulus set comprised relatively long horizontals and short verticals, and relatively small when the stimulus set comprised short horizontals with long verticals. To the extent that stimulus context exerts orientation-specific effects on perceived length, it thereby modulates the degree to which verticals appear longer than physically equivalent horizontals: the horizontal–vertical illusion (HVI). Under various contextual conditions, the HVI was as small as 3% (horizontals had to be 3% greater than verticals to be perceived as equally long) and as great as 15%, equaling about 12% in a "neutral" context. In Experiment 6, subjects judged the absolute physical length of each stimulus, and the results indicated that stimulus context acted largely by decreasing perceived lengths. The results are consistent with the hypothesis that differential effects of context reflect a process of stimulus-specific perceptual attenuation.

Stimulus context is well known to affect psychophysical judgments, which can be highly sensitive to such factors as the range of the stimuli, how often they are presented, and the sequence of presentation (for review, see Poulton, 1989). Especially striking are various stimulus-specific or *differential* effects of context—that is, situations in which context exerts unequal effects on different subsets of stimuli. Consider the example of loudness (Marks, 1988; also Marks, 1992b, 1994, 1996; Marks & Warner, 1991): In one condition, subjects judged, by magnitude estimation, the loudness of a tonal series comprising 500-Hz signals of relatively low sound pressure level (SPL) and 2500-Hz signals of relatively high SPL. In another condition, the 500-Hz tones were high in SPL and the 2500-Hz tones low, with subsets of stimuli at both frequencies common to the two conditions. The judgments of loudness showed strong differential effects across the two conditions: A given 500-Hz tone might be judged louder than a given 2500-Hz tone in the first condition, but softer in the second condition. Further, the shifts in relative loudness were substantial, with a 500-Hz tone of 70 dB being judged as loud as a 2500-Hz tone of 73 dB in the first condition, but of 57 dB in the second (Marks, 1988, Experiment 3).

Analogous effects emerged in judgments of intensity in other sensory modalities, such as taste (Rankin & Marks, 1991, 1992) and olfaction (Rankin, 1993). Although wide-spread, differential context effects are not universal but depend on the particular stimulus dimensions that vary. For example, shifting the *durations* of stimuli at 500 Hz and 2500 Hz across conditions had no effect on magnitude estimates of duration, whereas shifting SPLs at the same frequencies strongly affected estimates of loudness (Marks, 1992b). Similarly, shifting the physical lengths of lines presented in different colors had no differential effect on judgments of relative perceived length (Marks, 1992b), whereas shifting the physical lengths of lines presented in different spatial orientations did (Potts, 1991).

The present investigation received its instigation from the work of Potts (1991). Following the procedure of Marks (1988), Potts had subjects give magnitude estimates of the lengths of horizontal and vertical line segments interspersed within a single session. As with loudness, the judgments of relative length depended systematically on the set of stimulus magnitudes: When the physical lengths of the horizontal and vertical stimuli were the same (Experiment 1a), verticals were judged 14% longer than horizontals—an example of the well-known horizontal–vertical illusion (HVI). But the verticals were judged as little as 4% longer than corresponding horizontals when the verticals took on greater physical values than the horizontals (a constant set of rectangles was also judged—Experiment 4) and as much as 52% longer when the horizontals took on greater physical values than the verticals (and only isolated line segments were presented—Experiment 1b).[1]

Potts's (1991) findings suggest that differential effects of stimulus context are not limited to "intensive" perceptual continua, such as loudness, taste, and odor in-

tensity, but may also characterize an "extensive" spatial continuum, namely perceived length. Further, her findings imply that the quantitative metric of visual space itself is malleable: Not only is visual space anisotropic—its scale differing along the horizontal and vertical axes (as implied by the HVI)—but also, the degree of anisotropy—the magnitude of the HVI—can itself depend on the set of stimulus values presented to the subjects. But Potts's results also raise unresolved issues about the existence of a "neutral" context, about the mechanism underlying differential context effects, and about the relation between differential context effects and the HVI.

First, is there a "neutral" condition of stimulus context in which it is possible to obtain a "pure" measure of the HVI? We suggest that there is: A condition is neutral, with respect to differential effects of stimulus context, when the sets of horizontal and vertical stimuli are perceptually equal. This assertion rests, in turn, on the assumption that differential context effects depend on the existence of a discrepancy between the average perceptual levels. Note, however, that perceptual equality need not entail physical equality. If, everything else being the same, verticals appear greater than physically equivalent horizontals, then presenting physically identical lengths in the two orientations would mean that the verticals would be perceived on average as greater than the horizontals, and this difference in mean perceived length in turn could induce a small differential context effect. Thus, a neutral contextual condition would consist of a stimulus set in which the horizontals were sufficiently greater in physical size than the verticals so as to offset the HVI. One goal of the present study was to quantify the difference between perceived horizontal and vertical lengths under such a condition.

Second, what mechanism underlies differential effects of context? In our view, differential context effects represent the result of a high-level process of stimulus-specific perceptual attenuation. Recent papers on loudness have referred to these effects by the term *recalibration* (Marks, 1994, 1996). At first, it was suspected that differential context effects might represent the outcome of complex numerical response biases in magnitude estimation (Marks, 1988). However, similar effects turn up with a wide range of nonnumerical psychophysical methods, including paired comparisons of stimuli (Marks, 1993, 1994) and paired comparisons of stimulus differences (Schneider & Parker, 1990), so attributing differential context effects to decisional processes or "biases" in quantitative response seems dubious. Later, it was suggested that differential effects in loudness might reflect psychological "contrast," and in particular the operation of multiple, stimulus-specific mechanisms akin to Helson's (1964) adaptation level. Thus, at the sound frequency where SPLs are low, the adaptation level would be low and loudness elevated; at the frequency where SPLs are high, the adaptation level would be high and loudness lowered. But further evidence suggested that differential effects in loudness resulted from the relatively high SPLs presented but not from the low SPLs

(Marks, 1993), an outcome that is more consistent with a mechanism of attenuation or "suppressive adaptation" than with psychological contrast. Moreover, at high SPLs, this attenuation spread over sound frequencies, much like so-called auditory fatigue (Marks, 1994). Thus, a second goal of this study was to determine whether differential context effects in the perception of length are consistent with the model of attenuation.

A third question addresses the relation between differential effects of context and the HVI. Although Potts's (1991) results showed clear differential effects of stimulus context, the size of the HVI obtained in many of her test conditions was surprisingly large—as great as 52% in terms of magnitude estimates. More typically, the size of the HVI lies between 5% and 10%, measured as the percentage difference between matching vertical and horizontal extents (see, e.g., Prinzmetal & Gettleman, 1993). The exact value depends, however, on factors such as the "frame" encompassing the target (Künnapas, 1955, 1957) and the inferred depth of the display (Collani, 1985; Girgus & Coren, 1975), as well as the configural properties of the stimulus. Potts used simple, isolated lines as stimuli, whereas most studies of the HVI have used horizontal and vertical lines arranged in an L-shaped configuration, in which the HVI may be smaller (some studies have used a T-shaped configuration, which complicates the matter because of the presence of a "bisection illusion"; Finger & Spelt, 1947; Künnapas, 1955; Tedford & Tudor, 1969). Teghtsoonian (1972) conducted one of the relatively few studies that have avoided spatial configuration by using isolated line segments; using a matching method, she obtained an illusion of 11%. The unusually large values of the HVI that Potts reported might conceivably have been a consequence of the use of magnitude estimation. Note in this regard that Verrillo and Irvin (1979) also used the method of magnitude estimation and reported a very large HVI (equal to 36%).

More recently, the present authors reported magnitude estimates of the perceived length of verticals and horizontals and found both very small differential effects of context and a very small HVI (Marks & Armstrong, 1996, Experiment 1). Although the size of the HVI, about 4%, was more in keeping with other findings, the design of Marks and Armstrong's experiments was somewhat complicated in that its goal was to assess any possible interaction between the HVI in vision and its cousin in haptic touch, namely the *radial–tangential effect* (RTE), which refers to the tendency for arm movements made radially to the trunk of the body to feel longer in extent than physically equivalent movement made tangentially. Thus, in Marks and Armstrong's experiment, one group of subjects received contextual sets of stimulus objects oriented either horizontally and vertically to vision, and another group received the same stimuli requiring tangential or radial movements of the arm. In both groups, sets of haptic and visual stimuli both preceded the contextual stimuli (baseline trials) and followed them (aftereffect trials). Measures of aftereffects showed the differential effects of stimulus context to be modality

specific. That is, presenting different contextual sets to the eye produced no haptic aftereffects, and presenting differential contextual sets haptically produced no visual aftereffects. However, the differential effect of context in vision was itself so small as to be absent in the measures of aftereffects.

In a second experiment, Marks and Armstrong (1996) used a paired-comparison procedure instead of magnitude estimation, interspersing haptic test stimuli within the series of visual contextual stimuli and interspersing visual test stimuli within the series of haptic contextual stimuli. The results, even more clearly than those of the first experiment, showed differential context effects to be independent in the two modalities. Still, the intercalation of stimuli from a different modality might have affected the size of the context effect, the size of the illusion, or both.

A third main goal of the present study was to confirm, using a direct-comparison method, that the perception of length of lines differing in orientation, and hence the size of the HVI, depends systematically on the set of stimulus levels presented in different orientations, and in doing so to quantify both the magnitude of the differential context effect and that of the HVI under various test conditions. To minimize effects of stimulus configuration, we presented isolated horizontal and vertical line segments sequentially. The first five experiments used a paired-comparison method similar to one used by Marks and Armstrong (1996, Experiment 2). In this sequential procedure, subjects saw one line segment at a time, so the stimuli had no intrinsic 2-D configuration. The sixth experiment tested the hypothesis that the differential effects of context on perceived length involve a suppressive process of attenuation. This was accomplished by asking subjects to judge the *physical* lengths of horizontal and vertical lines presented in different contextual sets.

EXPERIMENTS 1 AND 2
Baseline Measures

Experiments 1 and 2 provided initial measures of perceived length in the horizontal and vertical orientations using the sequential, paired-comparison method. In both experiments, we paired each of five possible line lengths vertically with each of five lengths horizontally, presented at different locations in the viewing surface. The two experiments differed in the degree of uncertainty concerning the location of each stimulus presentation, with the uncertainty being greater in Experiment 1 than in Experiment 2.

Method

Subjects. All subjects in this and the following experiments were 18–40 years old and were recruited from the Yale community. Eight women and 2 men participated in Experiment 1, and 8 women and 4 men participated in Experiment 2. Six of the subjects participated in both experiments.

Apparatus. In this and the subsequent experiments, stimuli were presented on a 13-in. Apple color monitor driven by a Macintosh SE/30 computer equipped with an Exceed video card. The

room was dark except for the illumination provided by a small 4-W incandescent light and by the light emitted from the monitor. Responses were registered when the subject clicked a computer mouse on an appropriate location on the screen. Resolution on the screen was 640×480 pixels. In all of the experiments, the frame of the monitor and the outer portion of the screen were masked by a 34.5×34.5 cm matte blackboard containing an 18.5×18.5 cm aperture at its center, through which a square central region of the screen was visible. In all experiments, the subjects sat approximately 40 cm from the monitor.

Stimuli and Procedure. Line segments were black against a white background. Each line was 1 pixel thick. From the five lengths in each orientation (5.7, 6.0, 6.4, 6.8, and 7.3 cm), we constructed 25 possible combinations of pairs. The location of each line on the screen was defined by its center, which always fell on the midpoint of one of the screen's four quadrants. In Experiment 1, the horizontal and vertical lines presented in a given trial always fell in different quadrants. Thus, there were 12 possible pairs of location. Each of the 25 combinations of length at the two orientations was presented once in each possible pair of 12 locations, making 300 trials in all.

To determine whether uncertainty in stimulus location affects relative perception of horizontals and verticals, in Experiment 2 the horizontal and vertical lines always fell in the same location on the screen, both within a given trial and across all trials in a given session for a single subject. The location varied from subject to subject, however, with equal numbers of subjects receiving stimuli in each of the four locations. These locations constituted the corners of an implicit 3×3 cm square, concentric with the center of the screen. Thus, in Experiment 2, each of the 25 combinations of lengths was presented 12 times at the same location, again making 300 trials in all.

In both experiments, every trial contained a horizontal line and a vertical line, each lasting 0.5 sec and presented successively, with an interstimulus interval of 0.5 sec. The screen was blank for 0.3 sec between the termination of the visual warning signal and the onset of the first stimulus, and for 0.3 sec between the termination of the second stimulus and the onset of the signal to the subject to respond. Subjects clicked on one of two screen locations to indicate whether the first or second of the two lines looked longer. In half of the trials for each stimulus combination, the horizontal line came first, and in the other half, the vertical came first. Every session began with three randomly chosen practice trials, which were not analyzed. In both experiments, the sequence of stimulus pairs was randomly ordered and different for each subject. A brief break was given during each session. [2]

Results and Discussion

Responses were first scored as the proportion of times that the vertical stimulus was judged longer than the horizontal stimulus, $p(v > h)$ when pooled over the 12 possible pairs of locations for each stimulus combination (Experiment 1) or across the 12 replicates (Experiment 2), after which the proportions were converted to z scores, $[z(v > h)]$. Because several of the proportions equaled 0.0 or 1.0, we followed Tukey's (1977) recommendation of first converting proportions into split scores (*ss*), according to the formula $ss = (n + 1/6)/(N + 1/3)$, then transforming *ss* into z scores, which were then pooled across subjects.

The left panel in Figure 1 displays the family of functions obtained from Experiment 1, plotting the mean z score against the length of the vertical line segment, L_v; each of the five functions represents a fixed length of the horizontal segment, L_h. The right panel in Figure 1 dis-
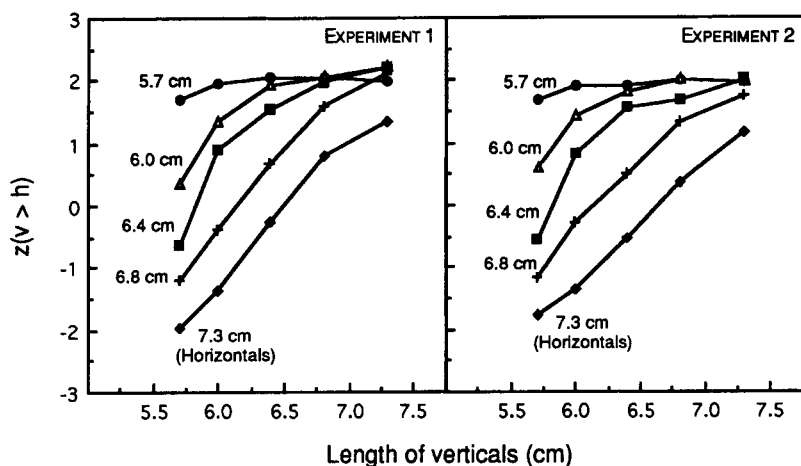
Figure 1. Probability that the vertical lines were judged longer than the horizontals, as a function of the physical length of the vertical. Each function represents a fixed level of horizontal length. The left-hand panel gives the results of Experiment 1, in which the location of the stimuli on the screen varied randomly over four possible values. The right-hand panel gives the results of Experiment 2, in which the stimuli appeared at a constant location for a given subject.

plays the corresponding functions obtained from Experiment 2. Were there no HVI, $z(v > h)$ should equal zero [the point of subjective equality, where $p(v > h) = .5$] when $L_h = L_v$. The fact that the actual values of $z(v > h)$ fall well above zero at these levels, on average, by about 1.5 normal deviates in both sets of functions, denotes the tendency for verticals to appear longer than physically corresponding horizontals.

On inspection, the values of $z(v > h)$ obtained at different levels of L_h within each experiment appear to define similarly shaped psychometric functions. This observation, together with Potts's (1991) evidence that the magnitude of the HVI was constant when expressed as a percentage, suggests that $z(v > h)$ should depend explicitly on the ratio of vertical to horizontal length, L_v/L_h, rather than on the absolute lengths. Consequently, we adopted this rule of proportionality in order to fit a set of lines to the entire family of functions, calculating a linear regression between all values of $z(v > h)$ and the ratio of L_v/L_h, separately for the data of Experiments 1 and 2. For this calculation, we used only z scores between +1.5 and −1.5 in order to avoid the residual end effects evident in Figure 1. The resulting regression equation for Experiment 1 had the form, $z' = -14.351 + 15.826 * L$ ($r^2 = .926$), and the equation for Experiment 2 had the form, $z' = -13.877 + 15.235 * L$ ($r^2 = .969$), where z' is the predicted z score and $L = L_v/L_h$. To compute the HVI, we set z' to zero in each case and solved for L, which turned out to equal 0.907 and 0.911, respectively. That is, at the point of subjective equality, a vertical line segment had to be only a little more than 90% as long as a horizontal segment. Alternatively, defining the percentage of illusion as $100 * [(L_h - L_v)/L_h]$, we infer that the resulting HVI was 9.32% in Experiment 1 and 8.91% in Experiment 2. The magnitude of illusion falls within the range reported in matching studies,

though somewhat on the high side, and is smaller than many of the values found by magnitude estimation.

Although the main characteristics just described seem to provide a more or less exhaustive account of the results, we also subjected both sets of z scores to an analysis of variance (ANOVA), in each case using the two within-subjects factors of horizontal length and vertical length, and incorporating the Huynh-Feldt correction for possible nonsphericity in repeated measures (in these and subsequent ANOVAs, we report corrected values of p). Both main effects were significant, in that the judgments decreased with greater horizontal length [$F(4,36) = 111.892$, $p < .0001$, in Experiment 1; $F(4,44) = 105.002$, $p < .0001$, in Experiment 2], and the judgments increased with greater vertical length [$F(4,36) = 161.922$, $p < .0001$, in Experiment 1; $F(4,44) = 85.725$, $p < .0001$, in Experiment 2]. In both experiments, the reliable interaction terms for horizontal length × vertical length [$F(16,144) = 16.779$, $p < .0001$, and $F(16,176) = 12.991$, $p < .0001$, respectively] largely reflected the end effects.

The results of the two experiments appear virtually identical. To confirm this inference statistically, we conducted an additional ANOVA on the data from both experiments using experiment as a between-subjects variable. There was no overall effect of experiment [$F(1,20) < 1$], nor was there any reliable interaction involving that term [experiment × horizontal length, $F(4,80) < 1$; experiment × vertical length, $F(4,80) = 1.546$, $p > .2$; experiment × horizontal length × vertical length, $F(16,320) < 1$], consistent with the inference that certainty or variability in location of stimuli on the screen had no effect of consequence on length perception. Together, the results suggest that the magnitude of the HVI, as measured with isolated line segments presented sequentially, is approximately 9%—a substantial value. Moreover, these measurements provide a baseline for examining how
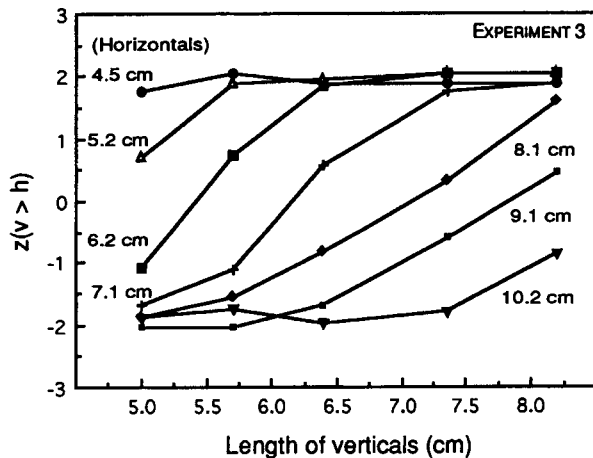
Figure 2. Probability that the vertical lines were judged longer than the horizontals, as a function of the physical length of the vertical. Each function represents a fixed level of horizontal length. In Experiment 3, the range of horizontal lengths encompassed the range of verticals, and mean horizontal length exceeded mean vertical length.

both length perception and the HVI depend on the context of stimulus levels presented in horizontal and vertical orientations.

EXPERIMENT 3
Differential Effects of Stimulus Set

The next three experiments probed the ways in which the stimulus context affects perceived length. The design of Experiments 1 and 2 led to a methodological concern and a theoretical consideration, both of which relate to the use of identical stimulus lengths at the two orientations. First, the methodological concern: Given the small number of lengths used, might the subjects have inferred, even implicitly, that the horizontals and verticals covered roughly equal stimulus ranges? If so, the results of the first two experiments might underestimate the "true" size of the HVI. In Experiment 3, we sought to reduce the likelihood of such an implicit inference by assigning different minimal and maximal lengths to the horizontal and vertical stimuli. Now, the range of vertical lengths fell wholly within that of the horizontals, the smallest horizontal being smaller than the smallest vertical and the longest horizontal longer than the longest vertical.

The second concern is theoretical. As noted, when horizontal and vertical lines are physically equal, the verticals will appear somewhat longer (the HVI). But because the verticals appear longer, the judgments are presumably subject to a differential context effect, which would act to diminish slightly the judgments of verticals relative to horizontals. In other words, the use of equivalent stimulus lengths in the two orientations should lead, ipso facto, to an underestimation of the size of the HVI. To circumvent this perturbation, a better paradigm

might present the subjects with *perceptually equivalent*, rather than physical equal, horizontal and vertical stimuli (see Marks et al., 1988, Experiment 9, for use of this same logic to assess the role of aging on odor perception). On the basis of the results of Experiments 1 and 2, we inferred that a "pure" HVI would slightly exceed 9%. Consequently, Experiment 3 sought to pre-equate the perceptual ranges of vertical and horizontal lengths by making the average length of the horizontal stimuli about 10% greater than the average length of the verticals.

Method
Subjects. Eight women and 4 men participated in Experiment 3. Nine had participated in one or both of the first two experiments.

Stimuli and Procedure. The apparatus and procedure were like those of Experiment 2, the main difference residing in the selection of stimuli, which consisted of seven horizontal lengths (4.8, 5.2, 6.2, 7.1, 8.1, 9.1, and 10.2 cm), each paired with each of five vertical lengths (5.0, 5.7, 6.4, 7.3, and 8.3 cm), making 35 stimulus pairs in all. Thus the set of horizontals is broader and the average length is greater than the corresponding values for the verticals (mean length = 7.24 cm for horizontals and 6.54 cm for verticals). Again, there were four possible locations on the screen, but only two were used for a given subject, balanced across subjects. Stimuli in each pair were presented either to the same location on the screen (half of the trials) or to the two different locations. Temporal order of vertical and horizontal stimuli in a pair was random. After three practice trials at the start of each session, which were not analyzed, the entire set of 35 possible stimulus pairs was presented a total of eight times, making 280 trials per session. Each session lasted about 30 min.

Results and Discussion
Data were analyzed as in the first two experiments; the mean z scores are displayed in Figure 2. An ANOVA performed on the z scores revealed reliable main effects of both horizontal length [$F(6,66) = 369.174, p < .0001$] and vertical length [$F(4,44) = 274.576, p < .0001$], as well as an interaction between them [$F(24,264) = 24.977, p < .0001$], largely signifying the end effects.

Again, the magnitude of illusion was determined by calculating a regression between $z(v > h)$ within the range -1.5 to $+1.5$, and the ratio $L = L_v/L_h$. The resulting equation was $z' = -10.322 + 11.804 * L$ ($r^2 = .923$). Once more solving for perceptual equality by setting $z(v > h) = 0$ gave $L = 0.8744$, and this in turn implies an illusion equal to 12.56%. That is, in Experiment 3, a horizontal line had to be more than 12% longer than a vertical line to appear equally long. This value compares to a difference of only about 9% in Experiments 1 and 2. The greater magnitude of the HVI in Experiment 3 was expected given the use of different minimal and maximal stimulus levels and especially given the patterns of differential context effects observed elsewhere. That is, the use of relatively greater horizontal than vertical lengths should make the horizontals appear relatively shorter than they would were the stimulus sets identical, and a reduction in perceived horizontal length corresponds to an increase in the size of the HVI. This is of
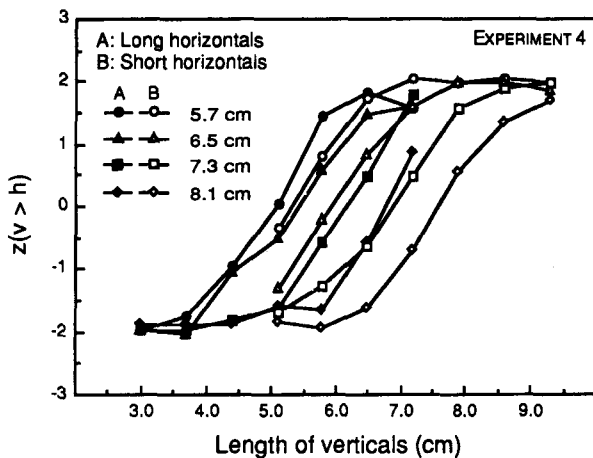
Figure 3. Probability that the vertical lines were judged longer than the horizontals, as a function of the physical length of the vertical. Each function represents a fixed level of one of the four horizontal lengths common to the two conditions. In Experiment 4, the overall set of horizontal lengths exceeded that of the verticals in order to make the perceptual ranges in the two orientations approximately equal. Stimuli were then presented in two contextual conditions. Condition A paired long horizontals with short verticals, and Condition B paired short horizontals with long verticals.

course to say that stimulus context affects the perception of length, or at least the judgment of length, and hence modulates the HVI.

Finally, note also that the slope of the mean psychometric function was only 11.8 in Experiment 3, compared with 15.5 in Experiments 1 and 2. The smaller slope in Experiment 3 is consistent with the use in that experiment of a greater range of horizontal stimuli compared with vertical stimuli, given the two assumptions. First, the slope of the psychometric functions is proportional to the ratio of the perceptual sensitivities at the two stimulus orientations. Second, sensitivity is inversely related to stimulus range. That is, given the stimulus reckoned as $L = L_v/L_h$, broadening the range of L_h relative to that of L_v should reduce the slope of the psychometric function.

EXPERIMENTS 4 AND 5
Symmetry and Asymmetry in
Differential Context Effects

Experiments 4 and 5 examined more explicitly how stimulus context affects length perception, and hence how context modulates the HVI. To do this, we adapted the varying-context paradigm of Potts (1991). In one condition, most of the verticals were short and most of the horizontals were long, and in another condition, most of the verticals were long and most of the horizontals were short. A subset of stimuli was common to each stimulus orientation in both conditions. Potts found the size of the HVI was greater when the verticals were short and the horizontals long. We followed Potts's method of varying stimulus contexts, but use the paired-comparison

procedure instead of magnitude estimation. To try to equate the overall sets of perceived vertical and horizontal lengths in Experiment 4, the average length of the horizontal stimuli exceeded that of the verticals. For comparison, in order to assess the role of the choice of mean stimulus lengths, Experiment 5 used vertical and horizontal stimuli chosen explicitly because they were mismatched in perceived length. In Experiment 5, the average physical length of the horizontals was smaller, rather than greater, than the average length of the verticals.

Method
Subjects. Seven women and 5 men participated in Experiment 4, and 6 women and 6 men participated in Experiment 5. Four subjects took part in both. Seven subjects in Experiment 4 had participated in one or more previous experiments, as had 10 subjects in Experiment 5.

Stimuli and Procedure. The apparatus and basic procedure were the same as those of Experiment 3. Ten horizontal and 10 vertical lengths were chosen for Experiment 4 in such a way that each horizontal length was 12% longer than the corresponding vertical. Horizontal lengths were 3.4, 4.1, 4.9, 5.7, 6.5, 7.3, 8.1, 8.8, 9.6, and 10.4 cm; vertical lengths were 3.0, 3.7, 4.4, 5.1, 5.8, 6.5, 7.2, 7.9, 8.6, and 9.3 cm. Different subsets of these stimuli were used in each of two stimulus conditions, with each condition pairing the seven shortest lengths in one orientation with the seven longest lengths in the other, making 49 stimulus pairs in all. Condition A paired long horizontals (5.7–10.4 cm) with short verticals (3.0–7.2 cm), and Condition B paired short horizontals (3.4–8.1 cm) with long verticals (5.1–9.3 cm). Thus, the four middle lengths at each orientation were common to the two conditions.

Locations of stimuli on the screen and procedures for randomization were the same as those of Experiment 3. Every session started with three practice trials, which were not analyzed. The entire set of 49 stimuli was presented in 8 replicates for a total of 392 trials per session, and each subject served in two sessions, held on different days, one with each stimulus condition. Sessions lasted about 40 min, with short breaks given one third and two thirds of the way through.

The procedure in Experiment 5 was identical to that of Experiment 4, except that the verticals were 12% greater, rather than smaller, than the horizontals. Thus the overall set of stimuli was strongly "biased" in Experiment 5 in terms of both the choice of stimuli and the HVI, so as to make the verticals appear considerably longer, on average, than the horizontals. Consequently, because differential effects of context should act to diminish the perceived vertical length relative to perceived horizontal length, we predicted that the HVI should be smaller in Experiment 5 than in Experiment 4.

Results and Discussion
The data obtained in each condition were analyzed as before, and the resulting z scores for the common horizontal stimuli appear in Figures 3 and 4 (Experiments 4 and 5, respectively). Most striking in each case is the displacement of the psychometric functions obtained in the two conditions. The data points corresponding to common stimulus values were higher in the context of Condition A, in which the other vertical lines were relatively short, than in Condition B, in which the verticals were relatively long. Thus, verticals were perceived to be relatively longer when the horizontals were long and the verticals short, and were perceived to be relatively shorter when the horizontals were short and the verticals long.
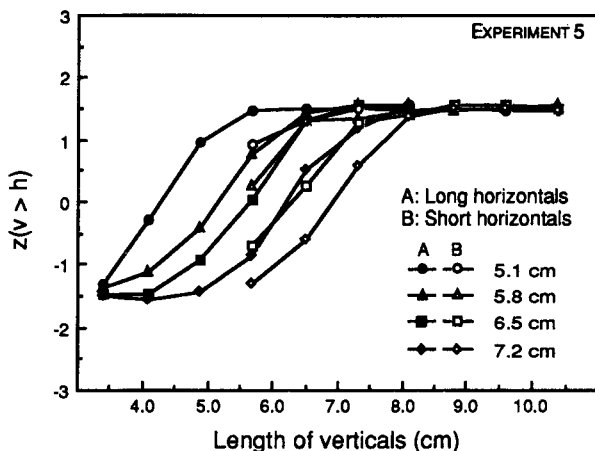
Figure 4. Probability that the vertical lines were judged longer than the horizontals, as a function of the physical length of the vertical. Each function represents a fixed level of one of the four horizontal lengths common to the two conditions. Again, Condition A paired long horizontals with short verticals, and Condition B paired short horizontals with long verticals. In Experiment 5, the overall set of vertical lengths exceeded that of the horizontals, so the perceptual ranges in the two orientations were mismatched.

To evaluate the results statistically, we performed, separately for each experiment, an ANOVA on z scores obtained for lengths common to Conditions A and B, using the between-subjects factor of order of sessions and the three within-subjects factors of contextual condition, horizontal length, and vertical length. First, consider results of Experiment 4. The ANOVA showed three reliable main effects: horizontal length [$F(3,30) = 160.538$, $p < .0001$], vertical length [$F(3,30) = 201.673, p < .0001$], and contextual condition [$F(1,10) = 103.237, p < .0001$]. The last term is crucial since it represents the differential effect of stimulus context. In addition, there were three reliable interactions: contextual condition × horizontal length [$F(3,30) = 6.290, p < .007$], horizontal length × vertical length [$F(9,90) = 8.333, p < .0001$], and contextual condition × horizontal length × vertical length [$F(9,90) = 6.571, p < .0001$]. These terms are difficult to interpret, since they seem to reflect, in part, compression of extreme probabilities (end effects).

Similar albeit not identical results were obtained in Experiment 5. Recall that the design was the same as in Experiment 4, but verticals were physically longer than horizontals. All four main effects were reliable. In addition to the main effects of horizontal length and vertical length [$F(3,30) = 121.984, p < .0001$, and $F(3,30) = 206.953, p < .0001$, respectively], there was a marginal effect of order [$F(1,10) = 5.074, p < .05$]. As with Experiment 4, however, the critical term was that of contextual condition, which again was reliable [$F(1,10) = 88.525, p < .0001$], once more reflecting the differential effect of context. Further, there were four reliable, but theoretically less consequential, interactions: contextual condition × horizontal length [$F(3,30) = 8.577, p < .002$], contextual condition × vertical length [$F(3,30) =

12.368, p < .0003$], horizontal length × vertical length [$F(9,90) = 29.260, p < .0001$], and contextual condition × horizontal length × vertical length [$F(9,90) = 3.531, p < .015$]. Again, end effects may have contributed to these interactions.

Once more, we regressed the z scores between $+1.5$ and -1.5 against the length ratio, L_v/L_h, in both experiments. Experiment 4 produced the following equations: $z' = -7.9387 + 9.3169 * L$ ($r^2 = .908$) for Condition A, and $z' = -9.056 + 9.7233 * L$ ($r^2 = .973$) for Condition B. Solving in each case for perceptual equality gave values of L of 0.8521 and 0.9314, respectively, or illusions of 14.79% in Condition A and 6.86% in Condition B. Experiment 5 produced the following equations: $z' = -8.7967 + 10.096 * L$ ($r^2 = .948$) for Condition A and $z' = -8.7677 + 9.0716 * L$ ($r^2 = .941$) for Condition B. Solving for perceptual equality gave values of L of 0.8713 and 0.9665, or illusions of 12.87% in Condition A and 3.35% in Condition B.

Experiments 4 and 5 showed HVIs in both stimulus conditions, although, as anticipated, illusions were greater in size in Condition A than in Condition B. And, again as expected, the illusion was correspondingly greater in each condition in Experiment 4 than in Experiment 5. In fact, the average illusion measured in Experiment 4, nearly 11% percent, was a little greater than the value of 9% found in Experiments 1 and 2, and more comparable to the value of 12% found in Experiment 3. In contrast, the average illusion in Experiment 5, about 8%, was the smallest of all.

We might chalk up these differences in the HVI to the selection of stimulus lengths. In Experiment 4, the horizontals were 12% longer overall than the verticals, whereas in Experiment 5 the verticals were 12% longer than the horizontals, and in Experiments 1 and 2 the lengths were identical. Thus the size of the HVI is directly related to the distribution of stimulus lengths. By our interpretation, the average illusions obtained in Experiments 3 and 4, in which the overall sets of horizontal and vertical stimuli were reasonably well "matched" for perceived length, presumably provide our best measures of the "pure" HVI.

EXPERIMENT 6
Assessment of "Absolute"
Changes in Perceived Length

To date, all of the studies of differential context effects have evaluated changes in relative suprathreshold response—that is, have evaluated changes in the magnitude of one kind of stimulus in relation to another kind of stimulus. In the domain of loudness, at least, evidence suggests that only relatively strong stimuli produce differential context effects (Marks, 1993, 1994). This finding is consistent with the hypothesis that differential context effects represent a kind of stimulus-specific suppression, produced by strong stimuli, rather than, say, a combination of suppression produced by the stronger stimuli in conjunction with enhancement produced by the weaker stimuli (as in an adaptation-level model; see
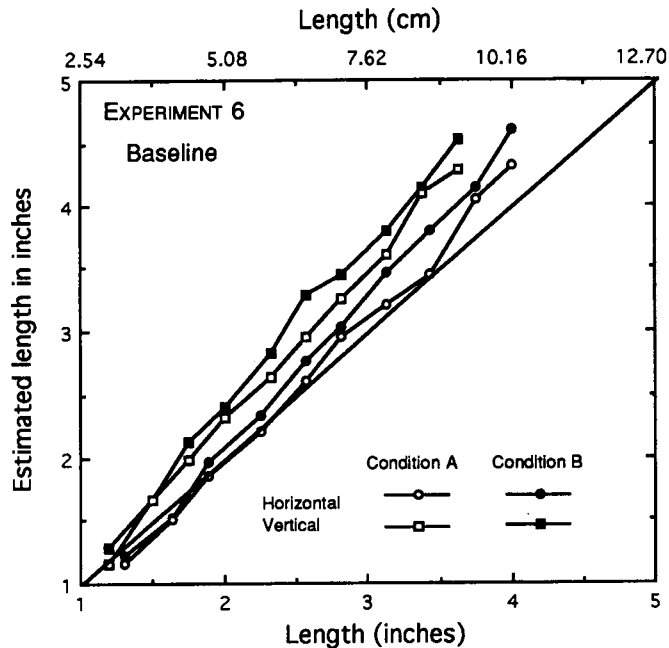
Figure 5. Baseline estimates of length, in inches (Experiment 6), of stimuli presented horizontally and vertically, prior to the differential context given in Conditions A and B.

Helson, 1964). It has been difficult to compare these two accounts experimentally, for to do so requires obtaining appropriate response measures on individual stimuli. Paired-comparison studies, of the kind used in Experiments 1–5, are intrinsically incapable of deciding, as they require relative judgments of stimuli presented in pairs. Methods such as magnitude estimation do provide measures on individual stimuli, but magnitude estimation has its own limitations, in particular, the possibility that the subject-defined or experimenter-defined number scale itself may change across stimulus conditions.

Relevant here is a study by Verrillo and Irvin (1979), who had subjects judge the length of isolated line segments using the method of "absolute magnitude estimation." Verticals and horizontals were presented in different test sessions, separated by at least 1 month, and from the results the authors inferred an exceptionally large HVI, amounting to 36% greater judgments of verticals. This estimate of the HVI seems overly large and may reflect in part an inappropriate assumption that the subjects' numerical scales were the same on the different occasions, a kind of invariance that has been called into question (e.g., Foley, Cross, & O'Reilly, 1990). A better strategy is to ask subjects to estimate lengths in physical units with which they are familiar and to make all of the judgments in a single session.

Masin and Vidotto (1983) used such an approach. They had subjects estimate line length in millimeters, without a standard, using, among other stimuli, isolated horizontal and vertical lines, and found a small HVI

(amounting to 3.9%). But Masin and Vidotto's study was not concerned with manipulations of the stimulus context.

In Experiment 6, we asked subjects to judge the lengths of horizontal and vertical lines in familiar physical units. After obtaining baseline measures on single horizontal and vertical lines, we presented pairs of stimuli, sequentially, in different contextual sets, as in the previous three experiments. Instead of merely asking subjects to indicate which stimulus in each pair was greater, however, Experiment 6 asked the subjects to judge the length of each. This paradigm made it possible, at least in principle, to determine whether perceived length declined when the lines presented at that orientation were large, whether perceived length increased when the lines presented at that orientation were small, or whether both kinds of shift took place. Further, this paradigm made it possible to determine local effects of context within each trial—that is, to determine whether the judgment of a given line's length was influenced by the length of the other line presented in the pair.

Method

Subjects. Eight women and 4 men participated. None had taken part in any of the previous experiments.

Stimuli and Procedure. The stimuli were 10 horizontal and 10 vertical line segments, the former being 10% longer than the latter in order to make the overall perceptual ranges approximately equal. Horizontals were 3.33, 4.14, 4.78, 5.72, 6.50, 7.14, 7.95, 8.74, 9.53, and 10.16 cm (1.31, 1.63, 1.88, 2.25, 2.56, 2.81, 3.13, 3.43, 3.75, and 4.00 in.). Verticals were 3.02, 3.81, 4.45, 5.08, 5.87, 6.50, 7.14, 7.95, 8.59, and 9.22 cm (1.19, 1.5, 1.75, 2.0, 2.31,
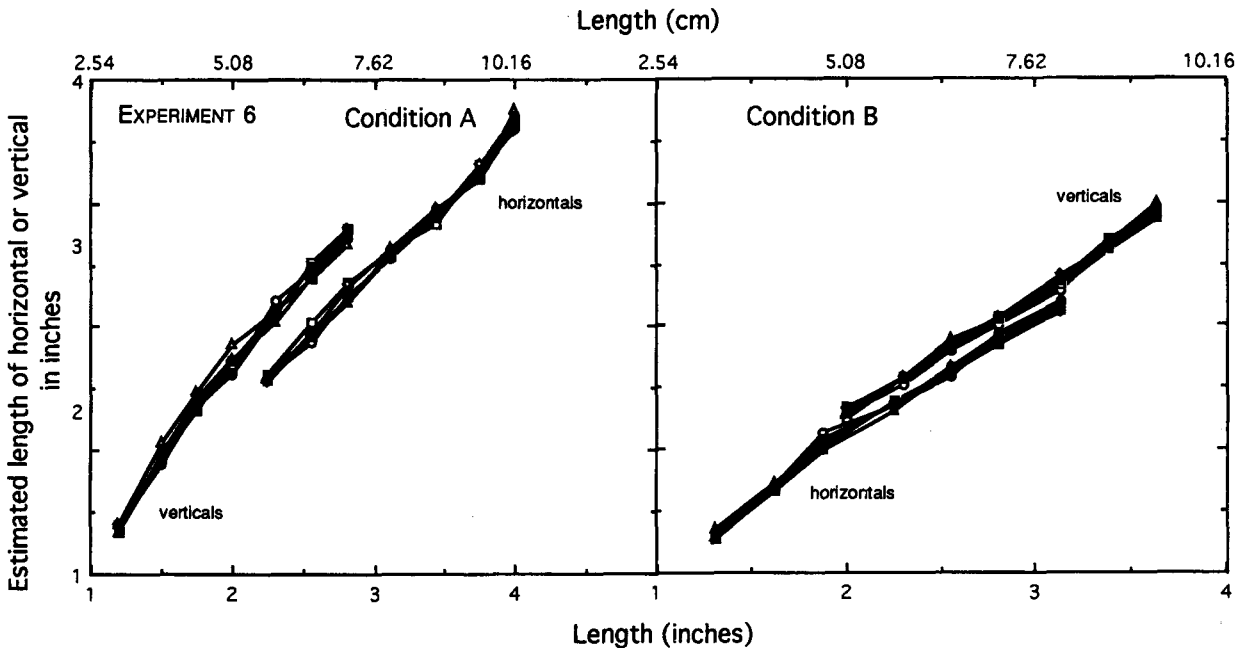
Figure 6. Estimates of lengths of lines oriented horizontally and vertically, as a function of their length, in two contextual conditions (Experiment 6). The overlap within each set of functions indicates that the level of the horizontal stimulus paired with each vertical stimulus had little effect on the judgment of the vertical, and that the level of the vertical stimulus paired with each horizontal stimulus had little effect on the judgment of the horizontal.

2.56, 2.81, 3.13, 3.38, and 3.63 in.). All stimuli were presented for baseline measurement, after which subsets of the horizontals and verticals were presented in the main part of each condition.

At baseline, after three practice trials, the 20 stimuli were presented for judgment one at a time in two replicates.[3] Subjects judged the length of each stimulus in inches (still the unit of length most familiar to young Americans), entering each response on a keyboard. After the baseline trials, the subjects received three practice pairs and then the main set of 49 pairs of stimuli in eight replicates. Stimulus pairs were formed by combining the seven longest horizontals with the seven shortest verticals in Condition A, and by combining the seven shortest horizontals with the seven longest verticals in Condition B. The main difference between this and the previous experiments was in the response. At the end of each trial, after the offset of the second stimulus of the pair, the subjects judged the lengths of both stimuli in the pair. Although this method requires the subject to encode and remember the length of the first stimulus in each pair while the second is being presented, it has the virtue of using the same temporal parameters as the previous experiments. Each subject served in two sessions, one for each condition, half participating first in Condition A and half in Condition B.

Results and Discussion

Baseline. Numerical estimates were averaged arithmetically across replicates within subjects and geometrically across subjects. The baseline measures, obtained prior to presenting the contextual stimuli of Conditions A and B, appear in Figure 5, in which the data are plotted against stimulus length in inches (scale at the bottom of the graph; here and in subsequent graphs, the corresponding scale of centimeters appears at the top of the graph).

The baselines reveal four main features. First, verticals were once again judged greater than physically equivalent horizontals, as evident in the displacement between each pair of vertical and horizontal functions. Second, although the horizontal lines were physically longer than the verticals, the average estimates given to horizontals and to verticals were nearly the same. On average, the verticals were judged 3% greater in length. Thus we were modestly although not completely successful in preselecting the stimulus levels to match perceptually. Third, the subjects tended to overestimate the verticals on their scale of "inches" but to judge the horizontals more nearly veridically. And fourth, the judgments obtained in the two sessions diverged a bit, with somewhat larger (5.5% greater) estimates given in Condition B than in Condition A.

The data (log transformed) were subjected to ANOVA, using the variables of condition, orientation, and length. Only two terms were significant—the main effects of length $[F(9,99) = 333.812, p < .0001]$ and orientation $[F(1,11) = 8.323, p < .015]$. Thus, the small difference in average measurements prior to Conditions A and B was not reliable. However, we were not wholly successful in matching horizontal and vertical lengths, since the verticals, although physically shorter than corresponding horizontals, were still judged slightly but reliably greater in length.

Contextual stimuli. Figure 6 shows the judgments of length given to horizontal and vertical stimuli in the main part of the experiment. Data from Conditions A and B appear in the left and right panels, respectively. Because each horizontal stimulus was paired with each vertical stimulus, the left-hand panel displays seven
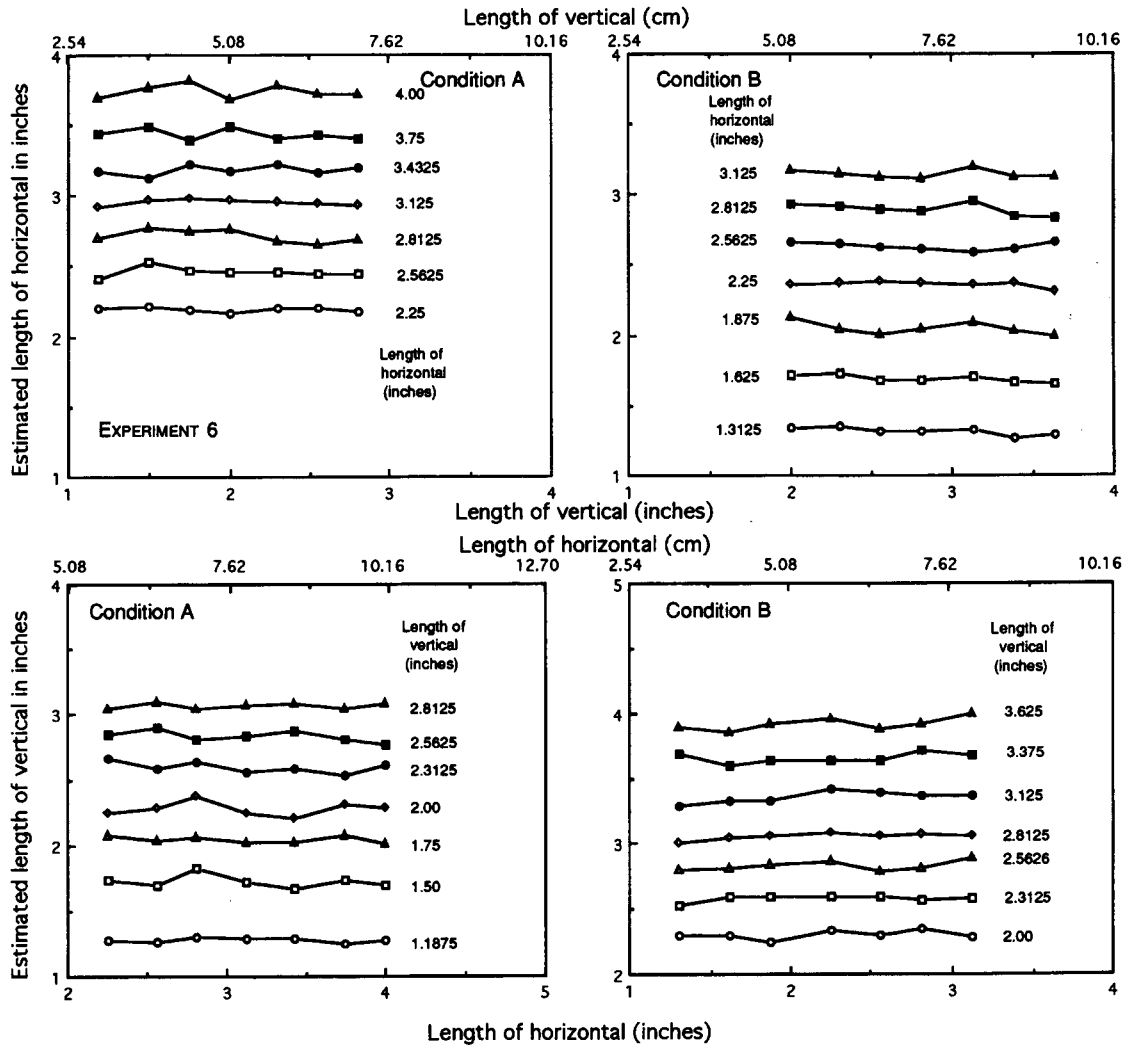
Figure 7. Data of Figure 6 replotted to show how the judgment of each horizontal stimulus varied as a function of the length of the vertical stimulus paired with it, and how the judgment of each vertical stimulus varied as a function of the length of the horizontal stimulus paired with it.

functions for horizontal judgments (one for each level of the vertical stimulus), and the right-hand panel displays seven functions for vertical judgments (one for each level of the horizontal stimulus). Two features of Figure 6 are notable. First, at a given physical length, verticals were judged substantially longer than horizontals in Condition A, but nearly the same as horizontals in Condition B. This is, once again, the differential effect of context. Second, the overlap among the seven functions within each set suggests considerable independence in the judgments at each orientation. That is, the judgments given to each stimulus in a given pair did not seem to depend substantially on the length of the other stimulus in the pair.

To examine in greater detail the influence of the "irrelevant" component of each stimulus pair, the data of Figure 6 are replotted in Figure 7. Judgments given to horizontals are plotted against the length of each vertical with which they were paired (upper panels), and judgments of verticals are plotted against the length of each

horizontal with which they were paired (lower panels). Data for Condition A appear on the left, and those for Condition B on the right. In general, the value of the paired stimulus seemed to exert very little effect, although there is a tendency in Condition B for the judgments of horizontals to decline slightly with increasing length of the verticals. To assess these effects statistically, the log-transformed judgments obtained in each condition for each orientation were subjected to ANOVA, using the variables of length of stimulus judged and length of paired stimulus. In each of the four analyses, the only reliable effect was that of the length of the stimulus judged [all values of $F(6,66) \geq 255.251$, $p < .0001$].

By limiting analyses to the data obtained with the four lengths at each orientation common to the two contextual conditions, it was possible to enter all of the data into a single ANOVA, using the variables of contextual condition, orientation, horizontal length, and vertical length. Five terms were significant. In addition to the main ef-
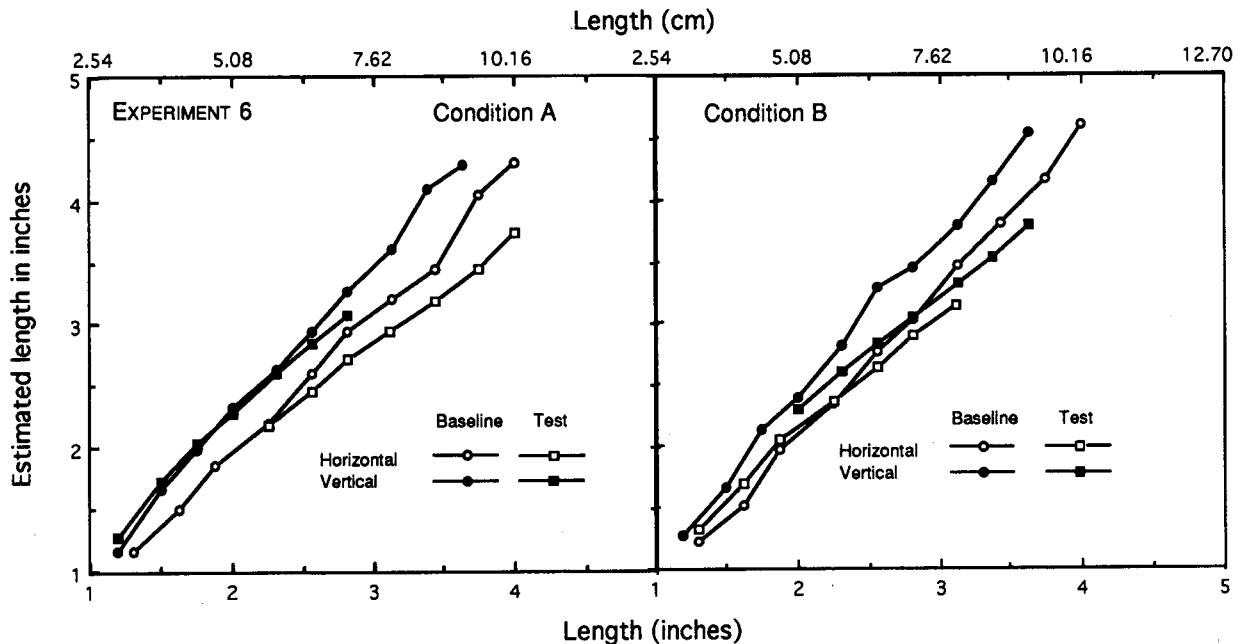
Figure 8. Averages of the judgments of length of horizontal and vertical stimuli obtained at baseline and in the two contextual (test) conditions in Experiment 6.

fects of horizontal length and vertical length [$F(3,33)$ = 160.537 and 92.356, respectively, $p < .0001$], there were three reliable interactions. The first was orientation × condition [$F(1,11)$ = 46.477, $p < .0001$], reflecting the differential effects of context. The other two, orientation × horizontal length [$F(3,33)$ = 267.874, $p < .0001$] and orientation × vertical length [$F(3,33)$ = 243.390, $p < .0001$], reflect the fact that the length of the horizontal stimulus affected the judgments of horizontals and verticals differently, and that the length of the vertical stimulus affected the judgments of verticals and horizontals differently. Further analyses on the stimuli common to the two conditions showed that judgments of horizontals depended on horizontal length [$F(3,33)$ = 324.007, $p < .00001$], but not vertical length [$F(3,33) < 1$], and that judgments of verticals depended on vertical length [$F(3,33)$ = 180.843, $p < .0001$], but not horizontal length [$F(3,33)$ = 1.068, $p > .35$].

Given the substantial degree of independence between the value of the stimulus judged and the level of the stimulus paired with it, we subsequently pooled data across the seven levels of the (irrelevant) paired stimulus to provide a single, average measure for each stimulus judged in each contextual condition. These overall averages appear in Figure 8 (squares) and are shown together with the corresponding baseline measurements (circles). Condition A appears on the left and Condition B on the right. In both conditions, the data indicate context-induced reductions in estimated length. In Condition A, in which the horizontals were long and the verticals short, the judgments of the horizontals declined from baseline,

whereas the judgments of the verticals were essentially unchanged. Similarly, in Condition B, in which the horizontals were short and the verticals long, the judgments of the verticals declined from baseline, whereas the judgments of the horizontals were largely the same.

These results were evaluated statistically by an additional ANOVA, this time using the variables of contextual condition (A or B), stimulus condition (baseline or test), orientation (horizontal or vertical), and length (four values). Five terms were significant or nearly so. The first two were main effects of length [$F(3,33)$ = 256.456, $p < .0001$] and, marginally, of orientation [$F(1,11)$ = 4.682, $p = .053$], the latter reflecting the tendency for the verticals to be judged as slightly longer than the horizontals despite our attempt to match them perceptually. The other three reliable terms were interactions of orientation × contextual condition [$F(1,11)$ = 5.591, $p < .04$], stimulus condition × length [$F(3,33)$ = 4.088, $p < .025$], and a three-way interaction of orientation × contextual condition × stimulus condition [$F(1,11)$ = 6.577, $p < .03$]. Post hoc tests (Newman–Keuls) showed that in Condition A, estimates of the length of horizontals decreased significantly from baseline ($p < .05$), whereas estimates of the verticals did not change. In Condition B, the estimates of the verticals decreased significantly from baseline ($p < .01$), whereas estimates of horizontals did not change.

In sum, these findings suggest that the main effect of stimulus context was a reduction in perceived length at the orientation in which the stimuli were greater. This outcome is consistent with our broader hypothesis that

differential effects of context result from a centrally mediated, adaptation-like process of perceptual attenuation.

GENERAL DISCUSSION

The perceived length of lines presented horizontally and vertically depended on the values of the stimuli presented in the two orientations. Relative to a baseline or "neutral" condition, increasing the physical lengths of the horizontals over those of the verticals led to an increase in perceived length of verticals, and increasing the physical lengths of the verticals over those of the horizontals led to an increase in perceived length of the horizontal. Thus, the metric of visual space is malleable, being subject to differential effects of stimulus context. Furthermore, these changes in perceived length appear to be largely, perhaps wholly, the result of decreases in length along whichever orientation the physical values are greater.

It is tempting to try to attribute these differential effects of context to some kind of decisional or judgmental process—for example, to a "relativity of semantic judgment" (see, e.g., Stevens, 1958), rather than to a change in underlying perceptual representations. We deem it unlikely, however, that an account in terms of decisional processes is correct. Several lines of evidence are hard to square with a decisional account. For example, it is not easy to see how such an account could explain why the judgments of length of lines depended differentially on the lengths presented vertically and horizontally, whereas comparable judgments of duration of lines did not depend differentially on the durations of exposure in horizontal and vertical orientations (Marks & Armstrong, 1996, Experiment 3), or why judgments of loudness of tones at different signal frequencies showed differential effects of context, but judgments of duration of tones at the same frequencies did not (Marks, 1992b). Although the existence of analogous effects in the judgment of loudness of low- and high-frequency tones was initially interpreted in terms of complex decisional processes (Marks, 1988), subsequent evidence suggested a more "sensory" or "perceptual" basis instead, perhaps some kind of central gain control (see Schneider & Parker, 1990) or adaptation-like attenuation (Marks, 1993, 1994).

An account in terms of stimulus-specific attenuation is readily compatible with the following four sets of findings, observed in various sense modalities:

1. Differential context effects are largely the same regardless of the psychophysical paradigm used to measure them, whether direct comparison (present experiments; Marks, 1994), direct comparison of differences (Schneider & Parker, 1990), or numerical estimation (Marks, 1988, 1994; Potts, 1991).

2. Effects of context on judgments of loudness depend mostly on relatively strong stimuli but not very much if at all on weaker ones (Marks, 1993), and effects of context on judgments of length appear mostly in the subset containing physically greater stimuli (present Experiment 6). Both of these results seem more compatible

with an underlying process of attenuation of strong signals than with, say, a Helsonian, adaptation-level process whereby all responses are shifted toward the middle of the subjective range, with strong signals being attenuated and weak ones amplified.

3. Judgments of loudness of tones differing in frequency show differential context effects when the signals fall in different critical bands (Marks, 1994; Marks & Warner, 1991), but do not show such effects (or at least show much less substantial ones) when the signals fall in the same critical band. Presumably, the mechanism underlying attenuation is specific to individual frequency channels.

4. Finally, and in a related vein, the presence of differential context effects seems to require that stimuli activate different neural subsystems, at least somewhat independently of their perceived similarity. Rankin (1993) found clear differential context effects when subjects sipped and judged different concentrations of sucrose solutions, which stimulate the gustatory system, and vanillin solutions, which stimulate the olfactory system, even though sucrose and vanillin are perceived as similar. Yet Rankin found the differential context effect to be absent when subjects sipped solutions that were perceptually similar—sucrose and sucrose-NaCl mixture, both of which stimulate only the gustatory system.

All of this evidence is consistent with the hypothesis that differential context effects reflect some kind of attenuation of responses in stimulus-specific sensory channels. Accordingly, this specificity should depend on the existence of distinct coding channels. Consider the case of duration. Duration is probably not encoded within channels specific to individual sound frequencies. As a consequence, judgments of duration have shown little or nothing in the way of differential context effects when the mean durations at different signal frequencies vary (Marks, 1992b). Loudness, however, is encoded in the auditory system within frequency-specific channels (critical bands). Presumably, attenuation can take place within individual, frequency-selective channels, being greatest in whichever channel the mean signal level is greatest. The difference in the magnitude of attenuation in different channels represents the differential effect of stimulus context. Note, however, that the frequency-selective channels in hearing arise largely in the auditory periphery. Yet differential context effects in hearing seem to rely substantially on a central neural mechanism—that is, at a higher neural level, since these effects show interaural transfer (Marks, 1996). Thus differential context effects in loudness appear to represent the central attenuation of information that initially derives from peripherally based channels.

Might analogous principles apply to the perception of length? If distinct channels encode length at different orientations, or at least in horizontal and vertical orientations, then the perception of length within these channels could be susceptible to stimulus-induced attenuation, which in turn would express itself as a differential effect of context. Pertinent here are the findings of Caelli

(1977), who had subjects compare the lengths of lines varying in shape ("squiggles," sinewaves) and who inferred from the results that the HVI is related to interactions between "orientation detectors" in the visual system. In Caelli's view, the perception of length is tied directly to mechanisms that underlie discrimination of stimulus orientation.

We suspect that differential effects of stimulus context operate at a level in the visual system beyond that of orientation detectors, much as the analogous effects in the perception of loudness arise in the auditory system beyond the level of the initiation of critical bands. In our view, the changes in perceived horizontal and visual lengths constitute changes in the perceptual metric, that is, compressions and decompressions of visual space. If so, the attenuation produced along a given spatial axis, whether horizontal or vertical, should be evident over the entire range of possible visual stimuli, not just over the range of stimuli used to induce it. Most likely, these changes take place in retinotopic coordinates, not in "external space." For if the metric of perceived length were tied to distal rather than proximal stimuli, it should also depend on contextual distribution along other stimuli dimensions, such as wavelength composition. But it does not (Marks, 1992b).

Regardless of their locus of origin in the visual system, differential context effects serve to modulate the relative perception of stimuli oriented horizontally and vertically and, in doing so, serve to modulate the size of the HVI. Indeed, one might ask whether the HVI itself might, in whole or in part, be the outcome of differential context effects. It has long been known that the HVI is sensitive to the "frame" of the visual field around the target (Künnapas, 1955, 1957, 1959). Indeed, a strong case can be made that the tendency for verticals to appear longer than horizontals across a wide range of conditions reflects the intrinsic shape of the visual field, which is wider than it is high (e.g., Prinzmetal & Gettleman, 1993). The typical explanation is that length is perceived relative to this frame. A given vertical line occupies a greater proportion of the vertical field than a physically equivalent horizontal line occupies of the horizontal field, and thus the vertical line encroaches on the upper and lower borders of the visual field to a greater extent than the horizontal line does on the left and right borders.

There might, however, be another possible way to account for the dependence of the HVI on the shape of the visual field, namely in terms of differential context effects. Because the visual field is wider than it is high, people may tend to experience greater horizontal than vertical extents. With binocular viewing, the visual field is ovoid, its horizontal axis being nearly half again greater than its vertical axis, about 200° versus 130° (see Prinzmetal & Gettleman, 1993). If, as a result of this asymmetry, people are exposed on average to greater horizontal than vertical extents, the long-term discrepancy

in the distribution of horizontal and vertical percepts might induce a differential effect on the perception of verticals and horizontals, enhancing the former relative to the latter and thereby producing the HVI.

In this regard, we take special note of the results obtained in Condition B of Experiment 5, in which the ratio of the average lengths of the horizontal to the vertical stimuli (all viewed within a square frame) was about 2:3. Thus, the mean stimulus levels were the inverse ratio of the axes of the natural, binocular visual frame, and might be expected to "cancel" the HVI. In fact, a small HVI of a little more than 3% remained. So differential context effects may not suffice to "explain" the HVI. Further, there is substantial evidence that the visual frame per se influences the perception of length of stimuli oriented in different axes. On the other hand, it is notable that the HVI persists even when, as in the present experiments, the stimuli are presented within a frame that is equally wide and high. That is, the HVI persists even in the absence of a concurrent asymmetric visual frame. Perhaps, then, although the HVI originates in the asymmetric shape of the natural visual field (Künnapas, 1955, 1957; Prinzmetal & Gettleman, 1993), its persistence is due to the long-term differential context that framing produces.

REFERENCES

CAELLI, T. (1977). Is perceived length affected by interactions between orientation detectors? *Vision Research*, 17, 837-841.

COLLANI, G. VON (1985). The horizontal–vertical illusion in photographs of concrete scenes with and without depth information. *Perceptual & Motor Skills*, 61, 523-531.

FINGER, F. W., & SPELT, D. K. (1947). The illustration of the horizontal–vertical illusion. *Journal of Experimental Psychology*, 37, 243-250.

FOLEY, H. J., CROSS, D. V., & O'REILLY, J. A. (1990). Pervasiveness and magnitude of context effects: Evidence for the relativity of absolute magnitude estimation. *Perception & Psychophysics*, 48, 551-558.

GIRGUS, J. S., & COREN, S. (1975). Depth cues and constancy scaling in the horizontal–vertical illusion: The bisection error. *Canadian Journal of Psychology*, 29, 59-65.

HELSON, H. (1964). *Adaptation level theory: An experimental and systematic approach to behavior*. New York: Harper & Row.

KÜNNAPAS, T. M. (1955). An analysis of the "vertical–horizontal illusion." *Journal of Experimental Psychology*, 49, 134-140.

KÜNNAPAS, T. M. (1957). Vertical–horizontal illusion and surrounding field. *Acta Psychologica*, 13, 35-42.

KÜNNAPAS, T. M. (1959). The vertical–horizontal illusion in artificial visual fields. *Journal of Psychology*, 47, 41-48.

MARKS, L. E. (1988). Magnitude estimation and sensory matching. *Perception & Psychophysics*, 43, 511-525.

MARKS, L. E. (1992a). The contingency of perceptual processing: Context modifies equal-loudness relations. *Psychological Science*, 3, 285-291.

MARKS, L. E. (1992b). The slippery context effect in psychophysics: Intensive, extensive and qualitative continua. *Perception & Psychophysics*, 51, 187-198.

MARKS, L. E. (1993). Contextual processing of multidimensional and unidimensional auditory stimuli. *Journal of Experimental Psychology: Human Perception & Performance*, 19, 227-249.

MARKS, L. E. (1994). "Recalibrating" the auditory system: The perception of loudness. *Journal of Experimental Psychology: Human Perception & Performance*, 20, 382-396.

MARKS, L. E. (1996). Recalibrating the perception of loudness: Interaural transfer. *Journal of the Acoustical Society of America*, **100**, 473-480.

MARKS, L. E., & ARMSTRONG, L. (1996). Visual and haptic representations of space. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI* (pp. 263-287). Cambridge, MA: MIT Press.

MARKS, L. E., STEVENS, J. C., BARTOSHUK, L. M., GENT, J. F., RIFKIN, B., & STONE, V. K. (1988). Magnitude-matching: The measurement of taste and smell. *Chemical Senses*, **13**, 63-87.

MARKS, L. E., & WARNER, E. (1991). Slippery context effect and critical bands. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 986-996.

MASIN, S. C., & VIDOTTO, G. (1983). A magnitude estimation study of the inverted-T illusion. *Perception & Psychophysics*, **33**, 582-584.

POTTS, B. C. (1991). *The horizontal–vertical illusion: A confluence of configural, contextual, and framing factors*. Unpublished doctoral dissertation, Yale University.

POULTON, E. C. (1989). *Bias in quantifying judgments*. Hove, U.K.: Erlbaum.

PRINZMETAL, W., & GETTLEMAN, L. (1993). Horizontal–vertical illusion: One eye is better than two. *Perception & Psychophysics*, **53**, 81-88.

RANKIN, K. M. (1993). *Differential effects of context on intensity judgements of taste and smell: Perceptual change or judgemental bias?* Unpublished doctoral dissertation, Stockholm University.

RANKIN, K. M., & MARKS, L. E. (1991). Differential context effects in taste perception. *Chemical Senses*, **16**, 617-629.

RANKIN, K. M., & MARKS, L. E. (1992). Effects of context on sweet and bitter tastes: Unrelated to sensitivity to PROP (6-*n*-propylthiouracil). *Perception & Psychophysics*, **52**, 479-486.

SCHNEIDER, B., & PARKER, S. (1990). Does stimulus context affect loudness or only loudness judgments? *Perception & Psychophysics*, **48**, 409-418.

STEVENS, S. S. (1958). Adaptation-level vs. the relativity of judgment. *American Journal of Psychology*, **71**, 633-646.

TEDFORD, W. H., JR., & TUDOR, L. L. (1969). Tactual and visual illusions in the T-shaped figure. *Journal of Experimental Psychology*, **81**, 199-201.

TEGHTSOONIAN, M. (1972). Apparent length as a function of tilt does not depend on orientation of the standard. *Journal of Experimental Psychology*, **94**, 191-197.

TUKEY, J. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

VERRILLO, R. T., & IRVIN, G. (1979). Absolute magnitude estimation of line length as a function of orientation and contrast polarity. *Sensory Processes*, **3**, 261-274.

NOTES

1. Given that judgments of length are approximately proportional to physical length, these measures of the HVI, constituting a percentage change in numerical judgment, are directly comparable to traditional measures, which are calculated in terms of percentage differences between perceptually equal vertical and horizontal stimuli.

2. To compare the use of a sequential as opposed to simultaneous presentation, we ran an auxiliary experiment using each type of presentation in a different session. The experiment was run under conditions of normal (fluorescent) room illumination, without the matte black mask (viewing surface = 26 × 19.5 cm). With simultaneous presentation, the two line segments appeared at a corner of the screen (the same corner was assigned to that subject in the corresponding session using sequential presentation). Horizontal and vertical lines in each simultaneous pair were placed 3.5 cm from the edge of the aperture, forming an L-shape with a gap between the segments. All of the horizontal and vertical segments were centered on the same two points, separated by 7.5 cm, so the minimal distance between the two segments was always at least 3.0 cm. Each session contained 300 trials (12 replicates of the 25 stimulus pairs), and the same 16 subjects served in both sessions, counterbalanced for order and held on different days.

Results suggested little difference between successive and simultaneous presentation. An analysis of variance (ANOVA) was performed on z scores (see main text) using the between-subjects factor of order of conditions and the three within-subjects factors of presentation (sequential or simultaneous), horizontal length, and vertical length. Only the main effects of horizontal and vertical length were reliable [$F(4,56) = 1,178.038, p < .0001$, and $F(4,56) = 856.158, p < .0001$, respectively]. Although the size of the illusion was slightly greater with sequential presentation than with simultaneous presentation (7.25% vs. 6.59%), the main effect of presentation was not reliable [$F(1,14) < 1$]. Three interactions were reliable: order × horizontal length [$F(4,56) = 2.567, p < .05$], horizontal length × vertical length [$F(16,224) = 111.994, p < .0001$], and, most importantly, presentation × horizontal length × vertical length [$F(16,224) = 5.839, p < .0001$]. Although not easily characterized, this three-way interaction implies that the method of stimulus presentation did in fact have a small, albeit highly stimulus-specific, effect.

3. Although we would have preferred to use the paired-comparison procedure in the baseline measurements, to do so would have required a substantially larger number of baseline trials in order to pair each vertical stimulus with each horizontal stimulus, and it was deemed more important to keep the number of baseline trials as small as possible so as to minimize any aftereffects induced by those trials (see Marks, 1992a, 1993).