

Size scaling: Retinal or environmental frame of reference?

DAVID J. BENNETT and WILLIAM WARREN
Brown University, Providence, Rhode Island

Previous studies have reported that when subjects are presented with two forms in a same-different task, their reaction times increase with the size ratio. This suggests a “mental scaling” transformation analogous to the “mental rotation” used to compensate for differently oriented forms in similar tasks. However, since the stimuli were presented in isolation, retinal and environmental size ratios were confounded. The present study varied both retinal and environmental size ratios in a same-different simultaneous matching task. In one experiment, random forms were placed at different distances along a textured hallway. A second experiment varied the monocular size information: In one condition the forms were displayed in a textured hallway; a second condition added cast shadows; and a third condition displayed the forms against a frontal wall of indeterminate distance. The results suggest that environmental size is determined prior to mental scaling and form matching.

A number of studies have found that when subjects are presented with two forms in a same-different form matching task, reaction times increase with the ratio of their sizes, even though size is irrelevant to the task (Bundesen & Larsen, 1975; Bundesen, Larsen, & Farrell, 1981; Ellis, Allport, Humphreys, & Collis, 1989; Howard & Kerst, 1978; Jolicœur & Besner, 1987; Larsen, 1985; Larsen & Bundesen, 1978; Sekular & Nash, 1972; see also Cave & Kosslyn, 1989; Corcoran & Besner, 1975; and Posner & Mitchell, 1967). This suggests a “mental scaling” transformation analogous to the “mental rotation” used to compensate for differently oriented forms in similar tasks (Shepard & Metzler, 1971). Two studies (Bundesen et al., 1981; Sekular & Nash, 1972) found additive effects of differences of size and orientation, suggesting separate, sequential processes. Several studies have varied the absolute sizes of the forms, while keeping the size ratios the same, and found that reaction time depends close to linearly on size ratio, not on absolute size difference (Besner, 1983; Bundesen & Larsen, 1975; Bundesen et al., 1981).¹ Simultaneous presentation (e.g., Bundesen & Larsen, 1975) and successive presentation (e.g., Larsen & Bundesen, 1978) yield qualitatively similar results.

In all of the forgoing size scaling experiments, the stimuli were presented in isolation against a blank background. As a result, retinal size ratios (or, alternatively, angular size ratios) were confounded with environmental size ratios, and so it is unclear which controlled the rise in reaction time. This issue is important, for whether size is recovered in environmental coordinates prior to or after mental scaling and form matching places a basic constraint on the functional architecture of high-level vision (Cave, Pinker, Giorgi, Thomas, Heller, Wolfe, & Lin, 1994).

This is well illustrated by considering the Kosslyn, Flynn, Amsterdam, and Wang (1990) theory of high-level vision. According to Kosslyn et al., information about the depth, distance, and spatial structure of facing surfaces—roughly corresponding to Marr’s (1982) 2½-D sketch—is computed by specialized input modules and collected in a “visual buffer.” At this point in the processing stream, size is said to be coded in retinal coordinates (Kosslyn et al., 1990, pp. 217–219). This view is supported by appeal to well-known neurophysiological studies of the visual cortex that are believed to reveal a number of visual maps organized retinotopically. Kosslyn has elsewhere associated the deployment of mental images with patterns of activity in such an early visual buffer (Kosslyn, 1980, 1987), in keeping with evidence that visual cortex is shared by both vision and visual imagery (Farah, 1990; Finke, 1989; Kosslyn, 1987). Since mental scaling is thus expected to occur in a visual buffer in which size is coded in terms of visual angle, the theory predicts that reaction time will be determined by retinal size ratio (see also Cave et al., 1994).

However, there are reasons to expect the opposite result. To begin, it is information about environmental size that is needed to guide behavior if we wish to pick up objects, walk through openings, or determine where to sit

We have benefited from comments from Kathryn Spoehr, Peter Eimas, Eric Sklar, Hal Sedgwick, Kyle Cave, and Myron Braunstein. Thanks also to Emre Yilmaz from D.B. for helping him learn graphics programming. David Ascher’s analysis program was used to help analyze the data for Experiment 2. Experiments 1 and 2 were presented in posters at the 1992 and the 1993 meetings of the Association for Research in Vision and Ophthalmology, respectively. Part of this work was supported by a postdoctoral research grant to D.B. by the Institute for Research in Cognitive Science, University of Pennsylvania. Correspondence should be addressed to D. Bennett, Department of Cognitive and Linguistic Sciences, Box 1978, Brown University, Providence, RI 02912 (e-mail: david_bennett@brown.edu).

or step (Gibson, 1950; see also the General Discussion, below). Further, Figure 1 illustrates that there is an immediate phenomenological effect of information for environmental size (the effect is more striking in the actual displays, viewed monocularly from behind a reduction screen). However, this does not decide the issue, since mental scaling and comparison may precede this conscious awareness. In the present experiments, we examined the hypothesis that mental scaling reaction times reflect estimates of environmental size.

In this study, we manipulated monocular or pictorial information for environmental size. Size perception is a complex subject, and we here settle for a brief sketch that indicates the basic sources of information available in the stimuli used in this study (Sedgwick, 1986, provides a particularly useful review and discussion, given present purposes; see also Cutting & Vishton, 1995; Gillam, 1995; Hochberg, 1971; McKee & Smallman, 1998). Important differences between the stimuli used in the different experiments and conditions are noted in the Method sections.

On the traditional account of size perception, size is determined by combining estimates of distance and of visual angle (this is the “size–distance invariance hypothesis”). Distance to a point on a ground surface can be determined monocularly, in eye height units, if the ground surface is known or assumed to be planar, and texture gradient (e.g., of size) is calculated at the corresponding point in the visual field (Purdy, 1958; Sedgwick, 1980; see also Sedgwick, 1986; Stevens, 1979, 1981). Given a planar ground surface and an explicit or implicit horizon, distance to a point can also be determined, in eye height units, from the visual angle between the point and the horizon (Sedgwick, 1980; Stevens, 1979; for evidence implicating “slope of regard” in distance judgments, see Wallach & O’Leary, 1982; see also Sedgwick, 1986).

In addition, a visible or implicit horizon specifies size directly, independently of distance—again relative to eye height, assuming that objects are resting on a ground plane (Sedgwick, 1980, 1986). Thus, the horizon inter-

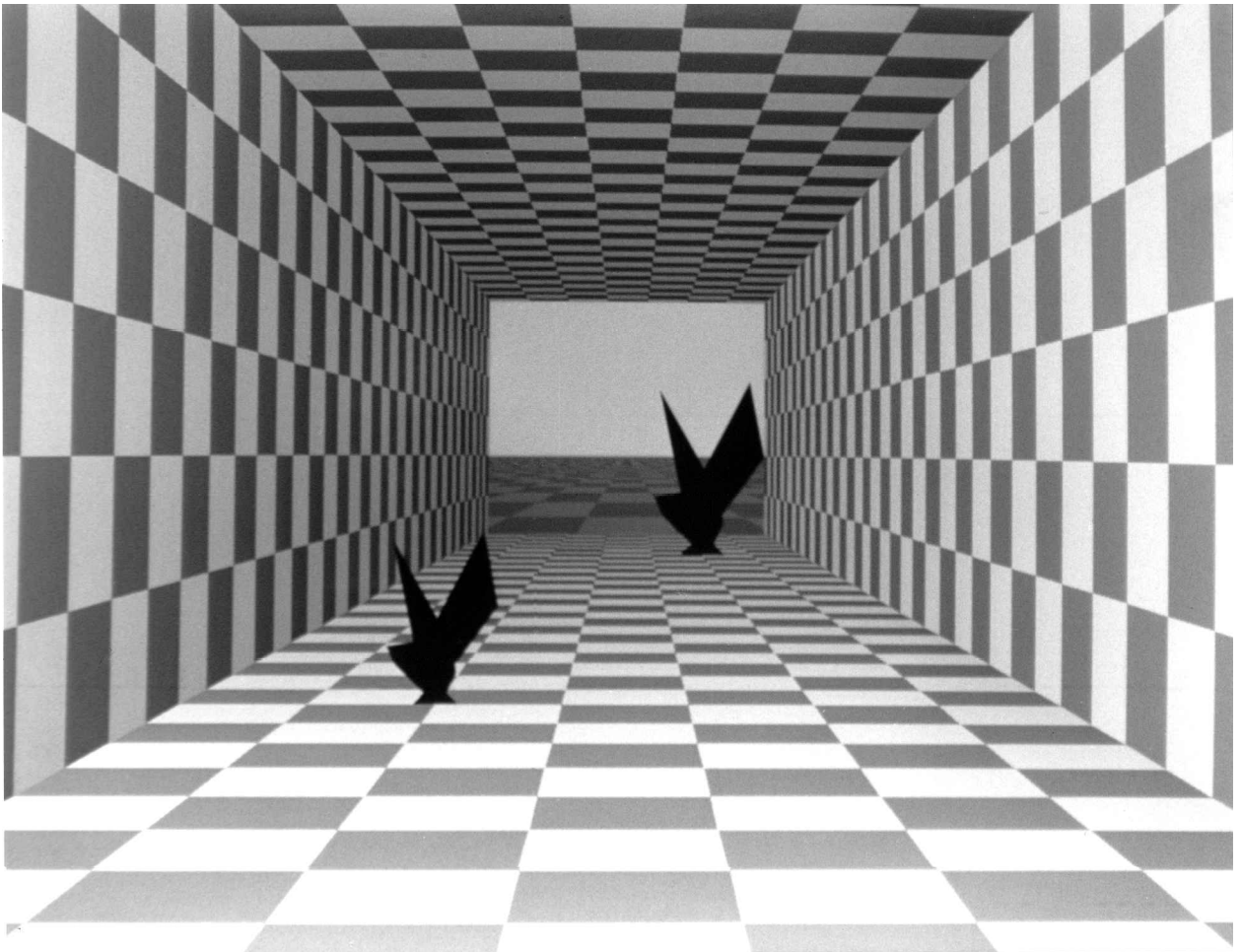


Figure 1. Black-and-white reproduction of an example stimulus (from Experiment 2, hall condition). The retinal size ratio is 1:1, but the environmental size ratio is 1:2.5.

sects objects taller than eye level at eye height. More generally, eye height scaled size along any dimension parallel to the picture plane can be determined provided that the visual angle between the horizon and the ground plane location directly beneath the surface being measured is available (Sedgwick, 1980; for empirical studies establishing and exploring the use of horizon-ratio information, see Mark, 1987; Warren & Whang, 1987; Wraga, 1999a, 1999b). Of course, if eye height is known in "absolute" units, such as "yards," then size can be determined in such units, even if initially "read-out" in terms of eye height.

It is, finally, possible that the texture elements might be used as a scene-intrinsic environmental unit of measurement, uninterpreted relative to body dimensions (or otherwise); the relative sizes of the forms are specified relative to the number of texture units that align with their bases. Similarly, making use of the display horizon, it is possible that the scene might be scaled up in terms of a viewpoint height unit, without giving a body scale meaning to this height (in effect making no commitment about the relation between the simulated ground surface and the real ground surface underfoot).

The present study was not designed to distinguish between differing theories of size perception. Rather, the aim was to explore whether size scaling is, in general, sensitive to information for environmental size, in order to address issues concerning the functional architecture of the visual system.

EXPERIMENT 1

The basic approach of the present study was to vary both the retinal and environmental size ratios in a simultaneous same-different form matching task by placing random forms at different distances along a textured hallway. For example, in Figure 1 the retinal size ratio is 1:1, but the environmental size ratio is 1:25 (as specified by the available monocular size information). A two-factor crossed design was used in which the two factors were environmental and retinal size ratios. There were five different levels of each factor, ranging from 1:1 to 1:3. The relative effect of retinal and environmental size ratio on reaction time can be determined by comparing the slopes associated with the main effects.

Method

Subjects. Nineteen subjects participated, 10 men and 9 women. Most were undergraduate or graduate students at Brown University, and all but 3 were paid \$10 for their participation. Six subjects had some general knowledge of the experimental design, including 3 who had run in an earlier pilot study; the remaining subjects were naive.

Apparatus. The displays were generated on a Silicon Graphics IRIS 4D/210GTX and displayed on a 19-in. monitor with a resolution of $1,280 \times 1,040$ pixels. They were viewed monocularly in a darkened room from behind a reduction screen with an aperture of 3.0×2.5 cm at a distance of about 19 in. A chin rest was used, but head movements were otherwise unconstrained.

Stimuli and Design. The checkerboard hallway was gray and white, the ground plane beyond was a reddish brown checkerboard,

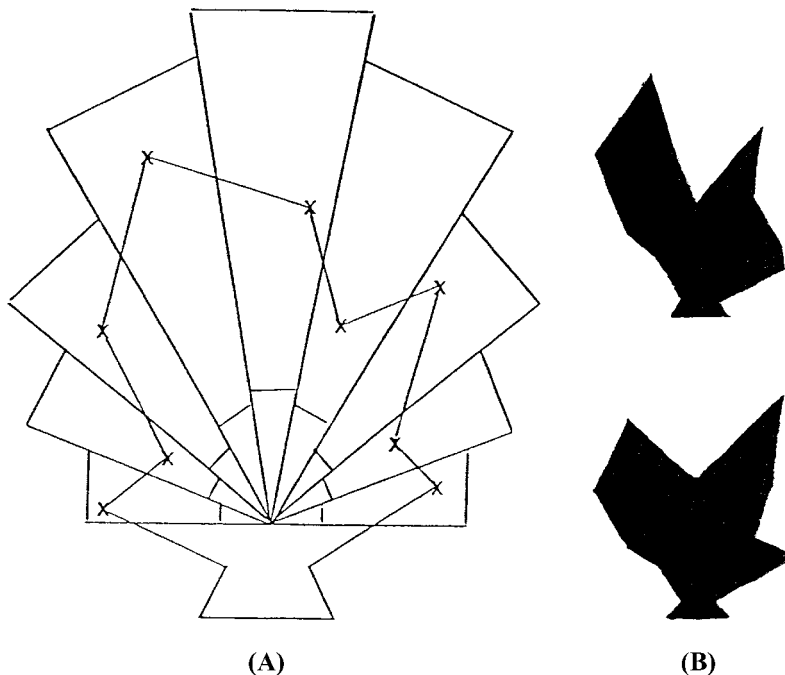


Figure 2. A: Construction of the forms (Experiment 1); B: example of different trial forms.

and the sky was blue. There was a slight darkening of the hallway and forms with distance. The forms were randomly varying silhouettes; nine vertices were chosen randomly within 20° “pie slices” and then connected together and placed on a base (Figure 2A). On same trials, the two forms had the same shape. On different trials, one of the forms was constructed by randomly perturbing the vertices of the other form (Figure 2B). The perturbations consisted of clockwise or counterclockwise rotations of 6.65° about the origin (the point at which the pie slices converge), and displacements in or out by 20% of the original distance from the origin; the directions of the perturbations were chosen randomly under the constraint that the vertices remain within the same region of the pie slice.

A two-factor crossed design was used, with five different retinal and environmental ratios: 1:1, 1:1.5, 1:2, 1:2.5, 1:3 (standard: variable). There were thus 25 different combinations of retinal and environmental ratios.

The *standard* form was always approximately .75 m high in environmental coordinates, assuming a standing eye height of 1.6 m above the floor. The *variable* form was generated and placed so that it conformed to one of the 25 combinations of ratios. There were 20 tokens of each combination for same trials and 20 tokens of each combination for different trials for a total of 1,000 trials. Same and different trials were randomly intermixed, as were tokens of each of the 25 combinations.

The placement of the standard form was determined as follows. The forms were placed within a range of 6–21 m from the station point in environmental coordinates. Notice that if the environmental and retinal ratios are the same, the forms are at the same distance down the hallway. If the retinal ratio (standard: variable) is less than the environmental ratio, then the standard form is closer than the environmentally larger, variable form—with the separation greatest when the retinal ratio is 1:1 and the environmental ratio is 1:3. The reverse is true if the retinal ratio is greater than the environmental ratio. So, for each of the 25 combinations of ratios, there was an allowable depth range in which the standard could appear. When the retinal and environmental ratios were the same, this was the entire 6–21 m; when they were not equal, this allowable range was smaller. For each of the combinations, this allowable range was broken up into five equally spaced intervals, and the standard appeared in each interval an equal number of times—with the interval and the exact location within the interval varied randomly.

The vertical visual angle of any individual form ranged from approximately 2° to approximately 20°.

Procedure. Each trial began with a black fixation bar (1.9°) shown sitting vertical in the middle of the hallway for 1.5 sec. Then a blank gray field was presented for 1 sec (to prevent apparent motion between the fixation bar and the forms), followed by the forms and the hallway. A pattern mask consisting of a grid of squares (each 1.5°) of randomly varying lightness was displayed as soon as a response was recorded. After the middle button was clicked, a blank gray field appeared for 500 msec, and then the next trial began.

Subjects were run in two sessions, each lasting approximately an hour. The sessions were run on different days within a week. The first session began with 130 practice trials, with feedback. This was followed by 500 trials, without feedback, in blocks of 100. The second session was exactly the same except that there were 100 practice trials. Subjects were instructed to ignore differences of size in judging whether the two forms were the same or different, and that “we will be analyzing how fast you respond on those trials where you answer correctly, so answer as quickly as possible while still remaining accurate.”

Results

The average overall error rate was 4.6%; incorrect responses were not included in the analysis. Outliers were

also eliminated according to the following rule: The mean and standard deviation were computed for each of the 25 combinations of size ratios for same trials, and outliers over three standard deviations were excluded for each combination. This always resulted in less than 2% of the trials being dropped.

The results are shown in Figures 3 and 4. All of the analysis of variance (ANOVA) *p* values reflect a Greenhouse–Geisser correction for within-subjects designs (Greenhouse & Geisser, 1959—the first method [of two] they present, as implemented in SPSS). A two-way repeated measures (RM) ANOVA on same trials revealed a significant effect of environmental size ratio [$F(4,72) = 24.644, p < .001$], indicating that the information specifying environmental size exerted a strong influence on reaction time. However, the effect of retinal size ratio was also significant [$F(4,72) = 14.489, p < .001$]. The environmental \times retinal interaction was not significant [$F(16,288) = 1.891, p = .093$].

T tests on individual subject slopes were conducted with a Bonferroni correction for the two tests performed, yielding an adjusted significance level of .025. The difference between the environmental slope (101 msec, Figure 3A) and retinal slope (71 msec, Figure 3B) was not significant [$t(18) = 1.705, p = .105$].

For those five combinations where the retinal and environmental ratios were the same, the two forms appeared in the same depth plane, as in previous size scaling experiments; these five combinations correspond to the diagonal five points shown in Figure 4. As shown in Figure 3C, the same depth plane slope of 184 msec was close to the sum of the environmental and retinal slopes (101 msec + 71 msec = 172 msec). A paired *t* test between the sums of the environmental and retinal slopes and the same-plane slopes was not significant [$t(18) < 1$].

The size ratio on the different trials may be defined in terms of the undistorted version of the form from which the perturbed version was generated; from trial to trial the size ratio on different trials might be slightly off, but the size ratios calculated this way will be roughly correct on average. Slopes on the different trials, thus defined, were flatter, suggesting a different process. However, the effect of retinal size ratio was still significant [$F(4,72) = 6.627, p < .001$], as was the effect of environmental size ratio [$F(4,72) = 4.087, p = .011$].

Error rates tended to increase slightly with increases in both environmental and retinal size ratio (Table 1).

Discussion

The results indicate that information specifying environmental size can strongly influence reaction time on a form matching task. However, there was also a strong and roughly equal effect of retinal size ratio. One plausible hypothesis consistent with the data is that, prior to the mental scaling and form matching, the visual system pools all available information and makes a “best estimate” about the environmental sizes or size ratio of the

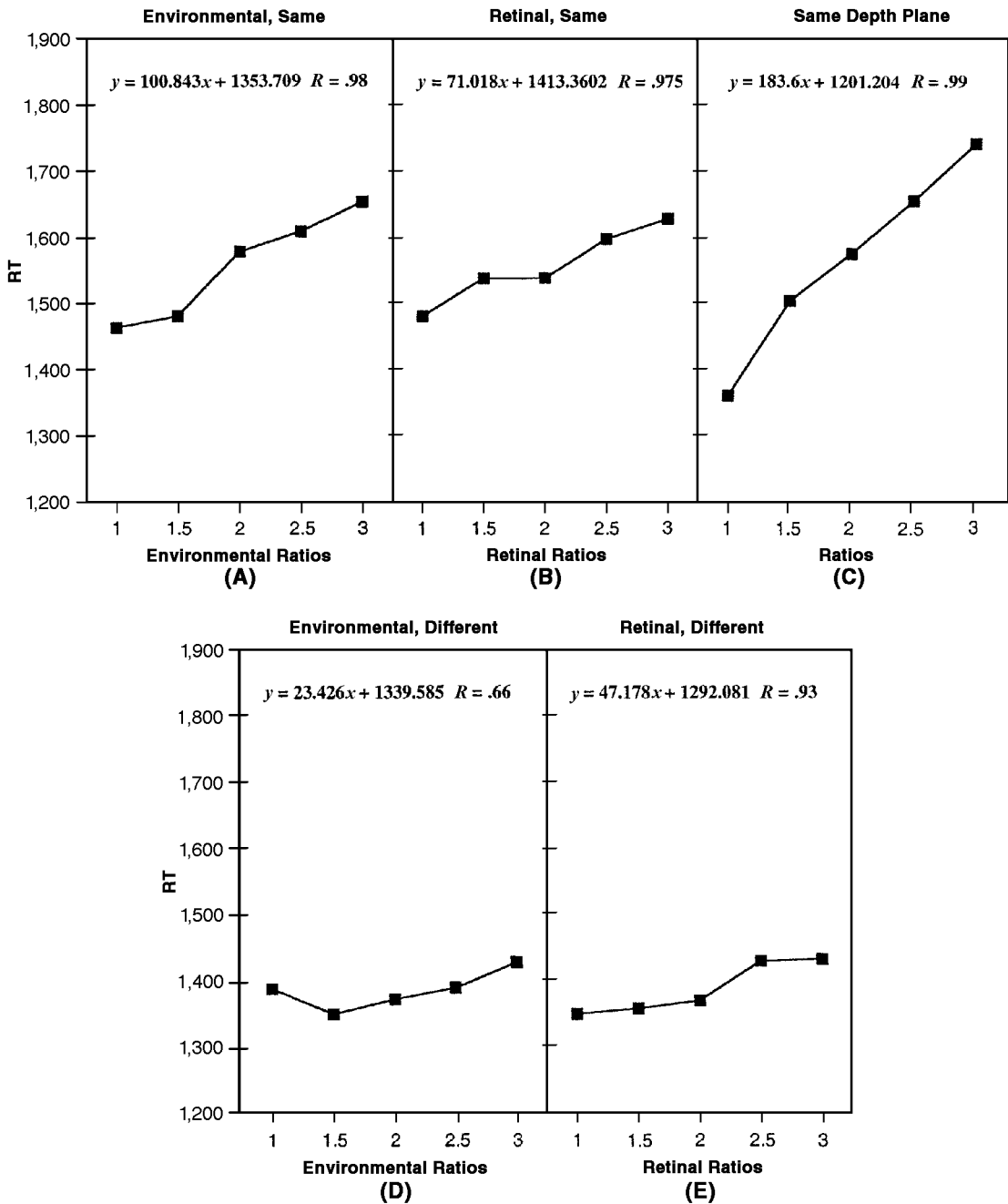


Figure 3. Experiment 1, (A) environmental size ratio, same trials; (B) retinal size ratio, same trials; (C) same-plane, same trials; (D) environmental size ratio, different trials; (E) retinal size ratio, different trials. To gain an intuitive understanding of these graphs, consider the rightmost data point of panel A: this data point represents the results of an equal number of trials of each of the five different retinal size ratios, with environmental size ratio remaining 1:3—so imagine two forms side by side that are 1:3 in both environmental retinal size ratio, and now gradually separate them in depth, passing through the other retinal ratios, until they are the same size retinally.

forms. On this hypothesis, the retinal size effect results from the conflicting and/or incomplete size information. For example, although the displays contained rich monocular size information, binocular disparity and

motion parallax were absent. Further, though care was taken to perceptually isolate the stimuli, some conflicting information that specified a flat screen remained: The viewing distance of 19 in. was close enough so that

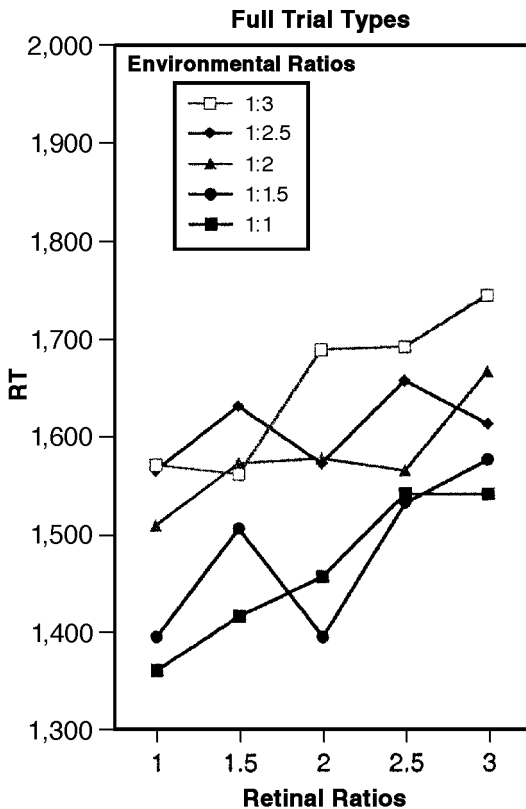


Figure 4. Experiment 1, full trial types, same trials.

accommodation may have had an effect; the edges of the screen may have occasionally come into view as subjects readjusted their heads; the same adjustments specify a flat screen by motion parallax. And to the degree to which the displays are interpreted as flat, the perceived size ratio of the forms will approach retinal size ratio—so “flat size ratio” is, on this account, a better phrase than “retinal size ratio.” When the environmental size ratio (as specified by the hallway information) and retinal size ratio disagree, the best estimate falls somewhere in between; when the forms are at the same distance, the retinal and environmental size ratios correspond, and the full size effect is obtained. This interpretation of the results is explored further in Experiment 2, described below.

Separation in depth. Separation of the forms in depth is not a confounding variable, since it is not correlated with retinal or environmental size ratio. But if separation in depth affected reaction time, this would lead to a systematic interaction of a specific kind. For example, consider the rightmost column of five data points in Figure 4. The upper right point corresponds to trials where the retinal and environmental size ratios are both 1:3, and so the two forms are at the same distance along the hallway. Moving down the column, the retinal ratio remains 1:3, while the environmental ratio gets smaller—and so the

forms grow farther apart. Thus, if there was an effect of separation in depth, the right sides of the bottom lines in Figure 4 should slope up relative to the right sides of the top lines. But there is no evidence of this pattern. Alternatively, if size and form information are recovered separately from distance information, then depth need not be involved in the form matching task.

Different trials. Performance on the different trials exhibited a quite different pattern from performance on the same trials. In same–different experiments, where there is a reaction time effect of a dimension irrelevant to the task (here, size) on same trials, performance on different trials is stimulus dependent, in ways that are not well understood. But flatter reaction times on different trials, like those observed in Experiment 1, are not uncommon (Farrell, 1985, especially pp. 429–330, 434–435). The present pattern of results suggests that subjects look for salient differences, which, if detected, allow them to quickly rule out the match; if no such salient difference is detected, there is enough uncertainty to lead them to scale one of the forms to check more carefully for a match—or, perhaps, to carry to completion a scaling process begun independently. The small but significant effects of size ratio on the different trials in Experiment 1 suggest that subjects do sometimes also scale when determining that the forms are different. These results diverge from those of Jolicœur and Besner (1987), who found a similar pattern for both same and different trials with a comparable range of size ratios (1:1–1:2.5) and different trials made up of different shapes (and not image plane rotations or mirror reflections). However, their forms had rectangular components which may not allow the sort of quick detection strategy sketched above, for reasons that are simply not clear. Unfortunately, the size scaling studies (Bundesen & Larsen, 1975; Larsen, 1985) that used stimuli most similar to those used in this study—spiky forms, randomly generated distractors, and a comparable range of ratios—do not report results for different trials. In the end, though, the issue does not seem to be of great importance, given current purposes: Subjects do scale on the same trials with these stimuli, and the question is whether size is coded in environmental or retinal coordinates when they do.

Further questions. An important question is whether the environmental and retinal slopes vary with the information available about the environmental sizes of the

Table 1
Same Trials, Error Rates

	Environmental Ratios					Retinal Ratios				
	1:1	1:1.5	1:2	1:2.5	1:3	1:1	1:1.5	1:2	1:2.5	1:3
Experiment 1	2.94	3.16	2.94	3.94	4.44	3.50	2.94	3.38	3.61	4.00
Experiment 2										
Hall	3.11	3.11	3.61	3.50	5.11	2.72	3.27	3.72	4.44	4.28
Wall	2.28	2.39	2.94	2.61	2.78	1.72	2.05	2.28	3.22	3.72
Shadows	2.83	3.11	4.11	5.44	5.11	3.44	2.94	4.61	4.78	4.83

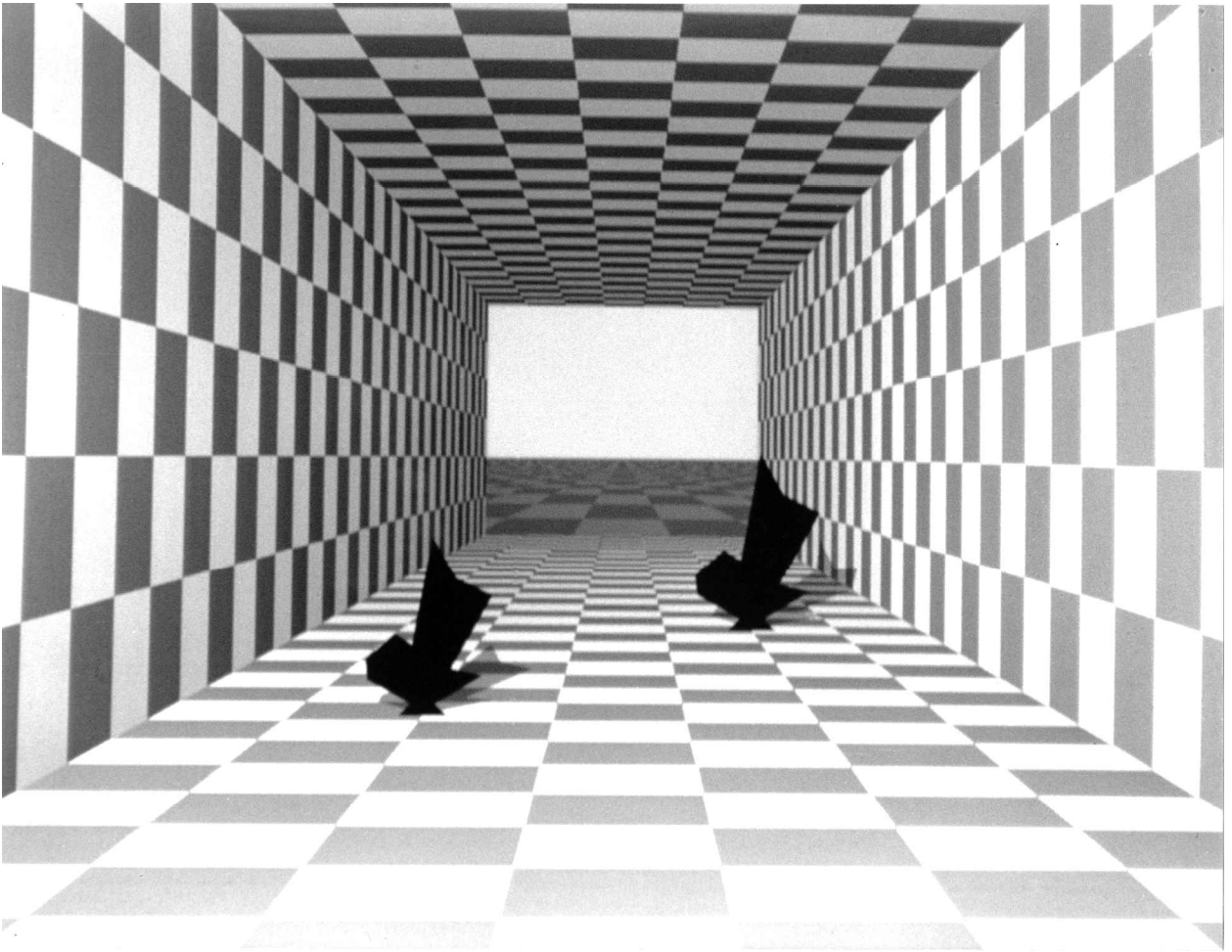


Figure 5. Experiment 2, example shadow stimulus.

forms. If the visual system is indeed pooling available information and taking a best guess in coding size prior to mental scaling, then the environmental size slope should increasingly dominate as the information specifying environmental size is enriched, while the retinal slope should increasingly dominate as environmental size information is removed.

The same depth-plane trials correspond to previous size scaling experiments, and the roughly linear rise in reaction time is qualitatively similar to the earlier size scaling results. However, because the same-different discrimination was fairly difficult, overall reaction times were higher than in previous size scaling experiments, which report reaction times of under a second for 1:1 size ratios (though the results are not out of line with reaction times seen in the mental rotation literature; see Folk & Luce, 1987). One possibility is that the longer the forms are displayed, the more time subjects have to somehow "take account of" the available environmental size information, thereby inflating the environmental size effect. These issues are addressed in Experiment 2.

EXPERIMENT 2

The sensitivity of the slopes to environmental size information was explored by varying the size information, within subjects, across three conditions.

The displays in one condition were similar to those used in Experiment 1, with the forms placed at different distances along a textured, slightly shaded hallway, with a salient horizon. A second condition reduced the distance and size information available by displaying the forms in the same screen locations as before but against a frontal wall. A number of influences on the perception of distance/size under reduced information conditions have been identified (see Sedgwick, 1986, for a review and discussion). In this condition, only the heights of the forms in the picture plane vary in accordance with levels of the environmental size ratio factor, and so would be the determining source of any remaining effect of "environmental" size ratio. Finally, a third condition added cast shadows to the hallway display (Figure 5). A possible explanation for the remaining, strong effect of

retinal size ratio observed in Experiment 1 was that subjects were able to partly ignore the size and distance information, “detaching” the forms from their surroundings. Other studies have shown that cast shadows strongly influence the perception of the location of static forms (Yonas, 1978), as well as the perception of the trajectory of motion in depth (Kersten, Mamassian, & Knill, 1997). So it was thought that adding cast shadows might help “anchor” the forms to the textured hallway, and thereby make it more difficult to ignore the environmental size information.

The reaction times in Experiment 1 were high because of the difficulty of the same–different discrimination. Although the low error rate shows that subjects were able to do the task, the high reaction times suggest that they did so only with difficulty. For Experiment 2, the same–different discrimination was made easier.

Method

Subjects. Twenty-one naive subjects began the experiment. Two dropped out after one session, one at the request of the experimenter after experiencing extreme difficulty with the task. A 3rd subject completed the experiment but was not included in the analysis because of an error rate that exceeded 20%. The remaining 18 subjects included 10 men and 8 women. Most were undergraduates at Brown University, and all were paid \$36 for their participation.

Apparatus. The apparatus was the same as that used in Experiment 1, except that a Silicon Graphics IRIS 40D/310 VGX was used instead of the GTX.

Stimuli and Design. The forms were generated in the same way as in Experiment 1, with three exceptions. First, the distortions on different trials were greater: Rotations were 6.65° as before, but the vertices were now perturbed in or out 25% of their original distance from the origin. Second, the forms were slightly narrower, to enhance the visibility of the shadows. Finally, the VGX polygon anti-aliasing was used to smooth the slight “jaggies” otherwise visible at the viewing distance used.²

Each of the three conditions in Experiment 2 had a separate, two-factor, crossed design, with five different retinal and environmental size ratios: 1:1, 1:1.5, 1:2.0, 1:2.5, 1:3.0. In the hall condition, the forms were displayed in a checkerboard hallway similar to that used in Experiment 1; the slight darkening with distance was now controlled by the IRIS real-time lighting model, with the simulated light source 3.2 m high, 1.6 m in front of the viewpoint, and 1.3 m to the left of the viewpoint. The shadows condition was formed by adding cast shadows to the hall condition, constructed using a shadow volumes algorithm (see Foley, van Dam, Feiner, & Hughes, 1990, chap. 16). The wall condition displayed the forms in the same screen locations as in the previous two conditions but against a frontal checkerboard wall of indeterminate distance.

Procedure. Each trial began with a fixation bar, shown in the middle of the hallway for 1 sec, followed by a blank gray field for 500 msec, and then the forms and the hallway. The pattern mask was displayed as soon as a response was recorded. After the middle button was clicked, a blank gray field appeared for 500 msec, and then the next trial began.

Subjects were run in six sessions on different days, each session lasting approximately 40–50 min. Each of the three conditions required two sessions to complete. All 18 subjects ran in all three conditions, with the order of conditions counterbalanced.³ The first session began with 110 practice trials, with feedback; remaining sessions began with 80 practice trials. For each session, the practice trials were followed by 500 test trials, without feedback, in blocks of 100. The instructions were the same as in Experiment 1.

Results

The results for the hall condition are shown in Figures 6 and 7. A two-way RM ANOVA on same trials revealed a significant effect of environmental size ratio [$F(4,68) = 21.671, p < .001$], indicating that the environmental size information still exerted a strong influence on reaction time after the changes in the stimuli. As before, the effect of retinal size ratio was also significant [$F(4,68) = 12.251, p < .001$]. There was also a small but significant environmental \times retinal interaction [$F(16, 272) = 2.327, p = .047$]. As in Experiment 1, the slope when the forms were shown in the same depth plane (142 msec) was close to the sum of the environmental and retinal slopes (75 msec + 61 msec = 136 msec). A paired t test between the sums of the environmental and retinal slopes and the same-plane slopes was not significant [$t(17) < 1$]. Slopes on the different trials were flatter, but environmental size ratio was still significant [$F(4,68) = 4.939, p = .008$], as was retinal size ratio [$F(4,68) = 3.207, p = .032$].

The results for the wall condition are shown in Figure 8. As expected, an ANOVA showed a strong effect of retinal size ratio in the absence of the hallway [$F(4,68) = 23.559, p < .001$]. However, the effect of the remaining environmental size information was also significant [$F(4,68) = 5.870, p = .006$]. The environmental \times retinal interaction was not significant [$F(16,272) = 1.522, p = .201$]. The slope of the forms shown in the same plane (138 msec) did not differ significantly from the sum of the environmental and retinal slopes (33.5 msec + 86 msec = 119.5 msec) [$t(17) = 1.845, p = .082$]. Once again, slopes on the different trials were flatter, but retinal size ratio was still significant [$F(4,68) = 4.812, p = .004$], as was environmental size ratio [$F(4,68) = 3.844, p = .01$].

The results for the shadow condition are shown in Figure 9. An ANOVA revealed main effects of both environmental size ratio [$F(4,68) = 9.093, p = .001$] and retinal size ratio [$F(4,68) = 16.529, p < .001$]. There was, in addition, a small but significant environmental \times retinal interaction [$F(16,272) = 2.612, p = .029$]. As in the previous condition and in Experiment 1, the same-plane slope (142 msec) did not differ significantly from the sum of the environmental and retinal slopes (59.5 msec + 65.5 msec = 125 msec) [$t(17) = 1.088, p = .292$]. Slopes on the different trials were once again flatter. Although retinal size ratio was still significant [$F(4,68) = 5.190, p = .009$], the effect of environmental size ratio was no longer significant [$F(4,68) = 1.393, p = .256$].

A condition (hall, wall, and shadows) \times effect (environmental and retinal) ANOVA on the individual subject slopes yielded a significant interaction [$F(2,34) = 15.893, p < .001$], in keeping with the dominance of retinal size when the hallway was removed. T tests on individual subject slopes were conducted with a Bonferroni correction for the 10 tests (including here the three comparisons concerning the same-plane trials already reported), yielding an adjusted significance level of .005. This is a conservative measure, and the p values are provided in case a weaker standard is felt appropriate.

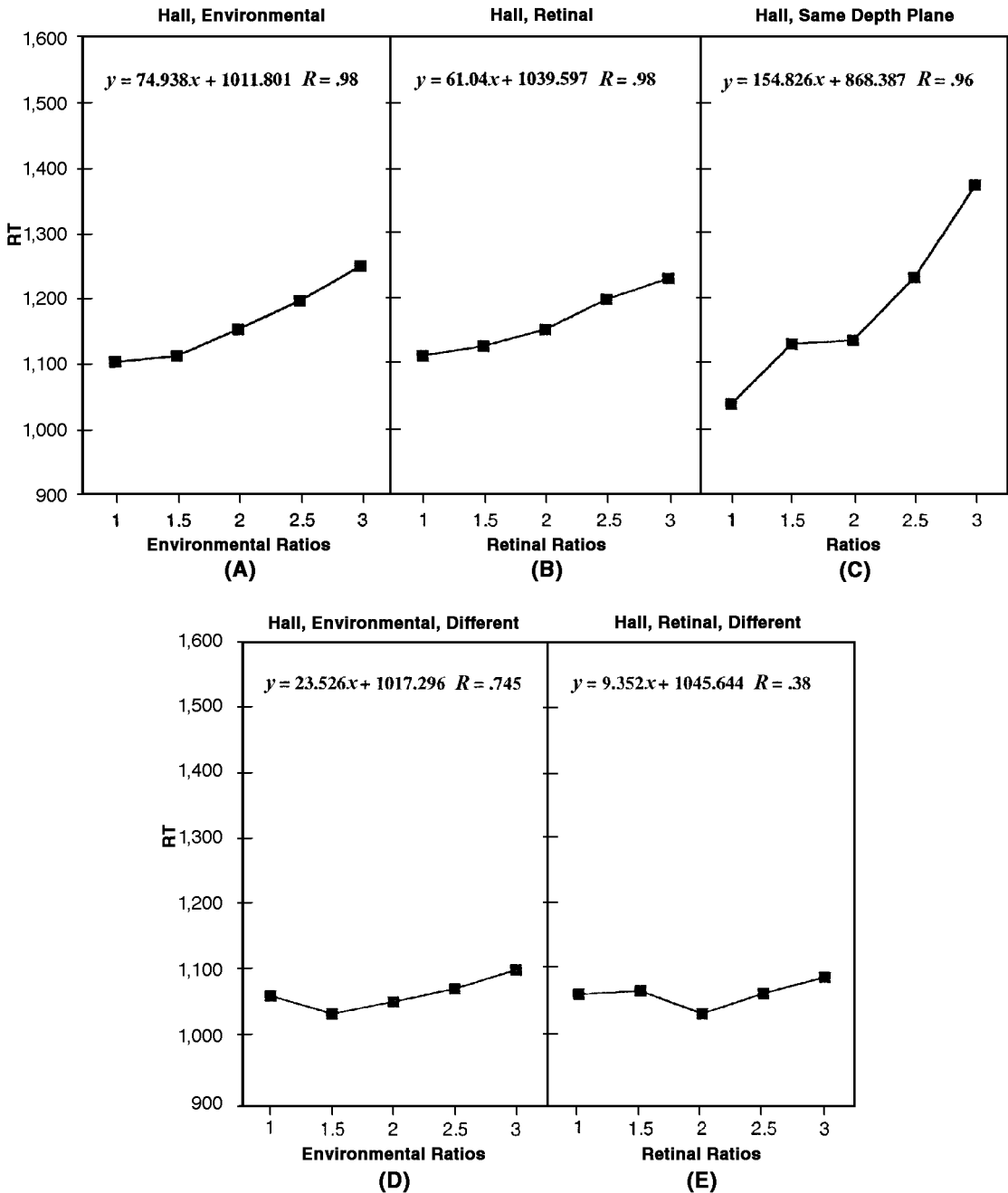


Figure 6. Experiment 2, hall, (A) environmental size ratio, same trials; (B) retinal size ratio, same trials; (C) same-plane, same trials; (D) environmental size ratio, different trials; (E) retinal size ratio, different trials.

In the hall condition, the difference between the environmental (75 msec) and retinal (61 msec) slopes was not significant [$t(17) = 1.22, p = .239$]. However, in the wall condition, the retinal slope (86 msec) was significantly greater than the environmental slope (33.5 msec) [$t(17) = 5.207, p < .001$], in keeping with the dominance of retinal size ratio in the absence of the hallway. The dif-

ference between the environmental slope in the hall condition (75 msec) and the environmental slope in the wall condition (33.5 msec) was also significant [$t(17) = 3.631, p = .002$], indicating that the full hallway information exerted a considerably stronger effect on reaction time than the weak size information that remained in the wall condition. On the other hand, the same comparison

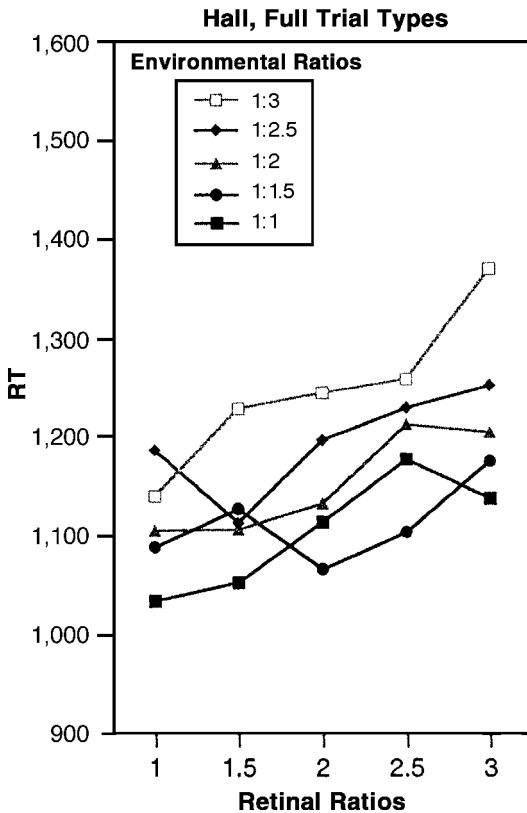


Figure 7. Experiment 2, hall, full trial types, same trials.

between the retinal slope in the hall condition (61 msec) and the retinal slope in the wall condition (86 msec) was not significant (after the Bonferroni correction) [$t(17) = 2.885, p = .01$]. The difference between the environmental (65.5 msec) and retinal (59.5 msec) slopes in the shadows condition was not significant [$t(17) < 1$]. Both the comparison between the shadows and the hall environmental slopes and the comparison between the shadows and the hall retinal slopes failed to reach significance, [both $t(17) < 1$]. This indicates the failure of the addition of the shadows to enhance the representation of environmental size.

The average overall error rates were 4.55%, 3.83%, and 5.105% in the hall, wall, and shadows conditions, respectively. Same-trial error rates tended to increase with both environmental and retinal size ratio (Table 1). A condition (hall, wall, and shadows) \times factor (environmental and retinal) \times ratio (1:1–1:3) ANOVA was conducted on the same-trial error rates. There was an effect of ratio [$F(4,68) = 10.26, p < .001$], reflecting the general rise in reaction time with retinal and environmental size ratios. Condition was also significant [$F(2,34) = 4.078, p = .035$]. However, there was no three-way interaction [$F(8,134) = 1.747, p = .134$]—so the different pattern of reaction times observed in the wall condition,

with the effect of retinal size ratio coming to dominate, was not due to a shift in speed versus accuracy strategy.

Discussion

The results from the hall condition were qualitatively similar to the pattern observed in Experiment 1. Thus, the strong effect of environmental size ratio remains with more discriminable stimuli and reaction times approaching those found in earlier size scaling experiments.

As expected, retinal (or “flat”) size ratio dominated in the wall condition. However, there was still a small but significant “environmental” size effect, due to information carried by the heights of the forms in the picture plane. In looking at the forms in the wall condition, there is indeed an impression that the higher of the two is more distant.

The failure to find a beneficial effect of shadows is perhaps, with the benefit of hindsight, not too surprising: Unlike the stimuli employed by Kersten et al. (1997) and by Yonas (1978), the forms used in Experiment 2 already possessed a perceptually salient base that clearly conveyed their location in the hallway. With such stimuli, cast shadows apparently do not provide additional useful location information. This is in keeping with Gibson’s (1950, pp.178–180) demonstration that there is a strong tendency to locate objects at the point they make optical contact with the ground plane, at least in the absence of evidence to the contrary.

GENERAL DISCUSSION

Best Distal Estimate or Proximal Mode?

The results of Experiments 1 and 2 are consistent with the hypothesis that, before the scaling and matching process is begun, observers determine a best estimate of environmental size, given all the available information. With the hallway conditions, the best estimate—when the “flat” and environmental ratios do not agree—is about halfway between that expected on the basis of the monocular scene environmental size information and that expected if the forms were at the same distance, on a flat surface such as the screen. In the wall condition of Experiment 2, the latter interpretation dominates. In all conditions, when the forms are at the same hallway distance, the retinal and environmental size ratios correspond and the full size effect is obtained.

However, in light of the remaining effect of retinal size ratio in the hallway conditions, there are two other interpretations of the pattern of results that must be considered.

One possibility is that the effect of retinal size ratio arises, at least in part, at a different level of visual processing than the effect of environmental size ratio. Specifically, Cave and Kosslyn (1989) provided evidence for a size-based process of selective attention that is independent of the mental scaling employed in making shape comparisons (see also Larsen & Bundesen, 1978). However, even if it is assumed that this attentional process is retinally coded and that subjects in Experiments 1

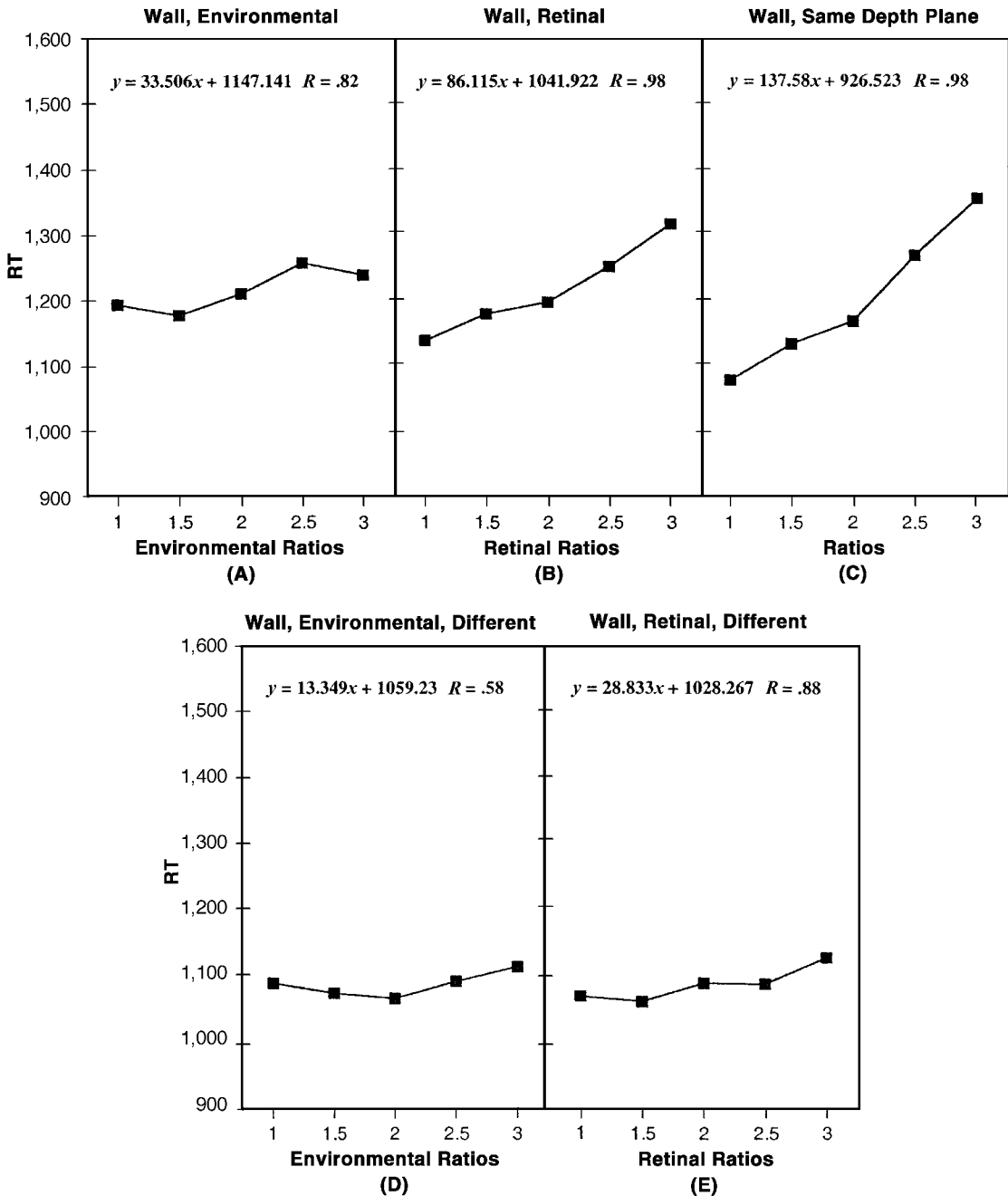


Figure 8. Experiment 2, wall, (A) environmental size ratio, same trials; (B), retinal size ratio, same trials; (C) same-plane, same trials; (D) environmental size ratio, different trials; (E) retinal size ratio, different trials.

and 2 regularly “resized” an attentional window (say, once per fixation), it is still unlikely this would account for all, or even much, of the retinal size ratio effect observed in Experiment 1 and in the hall and shadow conditions of Experiment 2. This is because of the small slopes associated with such attentional adjustment, approximately 4–9 msec—with the lower number probably the closer estimate of the speed of attentional resizing,

per se (see Cave & Kosslyn, 1989, who also discuss and compare Larsen & Bundesen, 1978). In any case, to the degree to which the retinal effect is traceable to attentional resizing, the conclusion that subjects code size environmentally prior to scaling would be strengthened.

Another possible explanation of the pattern of results is suggested by Rock’s (1983) observation that “although an object at varying distances does appear to be the same

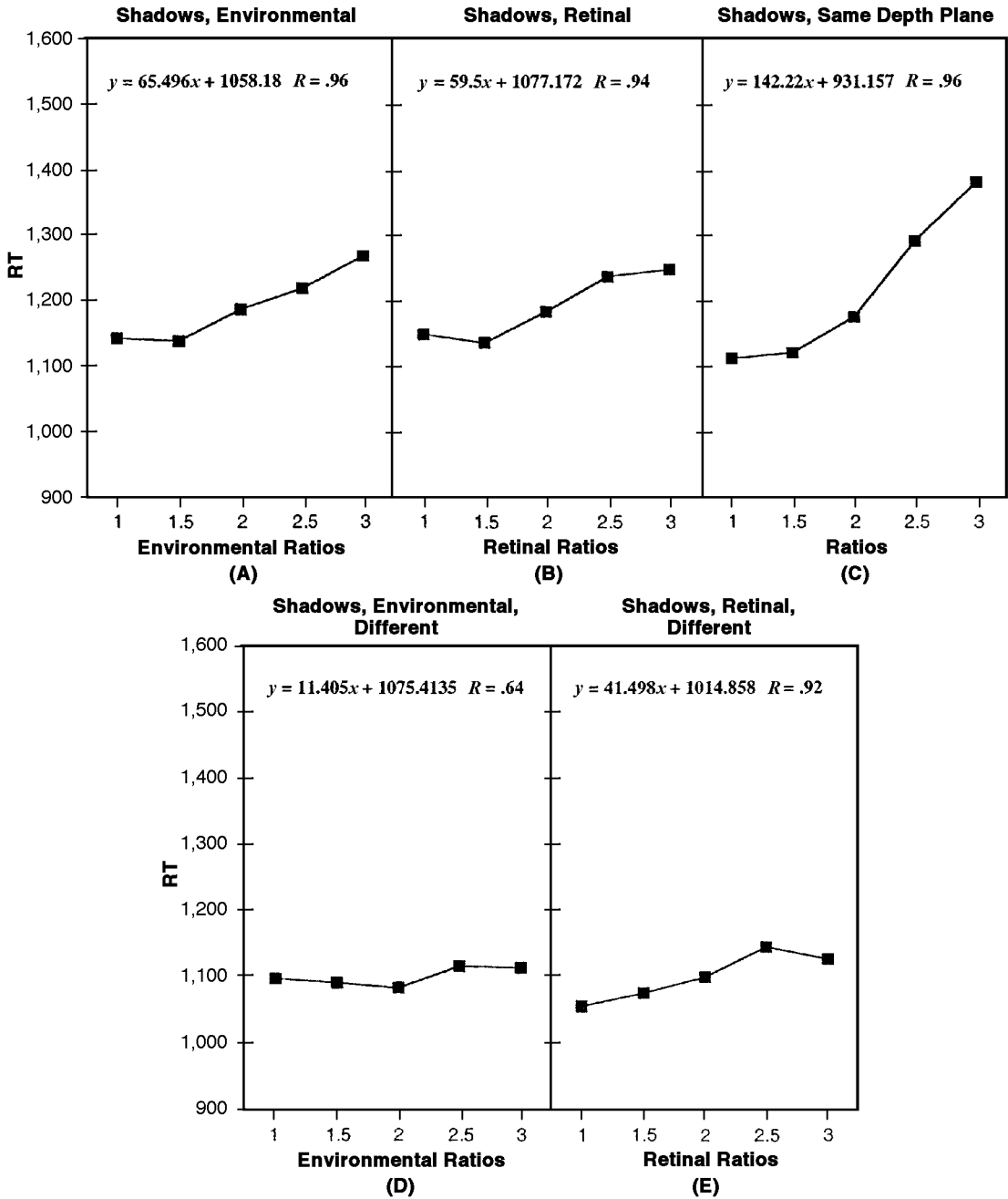


Figure 9. Experiment 2, shadows, (A) environmental size ratio, same trials; (B) retinal size ratio, same trials; (C) same-plane, same trials; (D) environmental size ratio, different trials; (E) retinal size ratio, different trials.

objective size, its changing visual angle is by no means without representation in consciousness. We are aware of, even if not attending to, the fact that at a greater distance the object does not fill as much of the visual field of view as it does when it is nearby” (p. 254). Rock called this the “proximal mode” of perception, in contrast to the more natural “world mode,” in which we attend to and represent environmental size. Gibson (1950, pp. 26–43) described a similar attitude of attending to the “visual

field,” which requires an “introspective or analytic” withdrawal from the more natural, everyday, immersion in the “visual world.”

It is doubtful that this is truly a *proximal* mode, since (as Rock acknowledges) when subjects are instructed to adjust a comparison object to match a test object according to visual angle, their settings are between those that would match visual angle and those that would match environmental size (typically about halfway—see

Sedgwick, 1986; see also McKee & Smallman, 1998, and McKee & Welch, 1992). The theoretical interpretation of the proximal mode, as well as its exact relation to size judgments rendered under various instructions, are difficult issues and matters of some debate (Carlsen, 1977; Epstein, 1963; Gogel, 1990; McKee & Smallman, 1998; Thouless, 1931). But Rock and Gibson are surely right that there is a perceptual stance or attitude that leads to judgments of size that at least more closely reflect visual angle than judgments rendered in the everyday, "world mode." Thus, an alternative explanation of the pattern of results observed in Experiments 1 and 2 is that subjects adopt the proximal mode, and as a result the representation of the size of the forms reflects a partial failure to use, or "tuning out," of the environmental size information in the displays.

Is there any way to empirically distinguish the possibility that subjects determine size in the "world mode"—but their representations reflect incomplete and conflicting size information—from the possibility that subjects are determining size in the proximal mode? One approach would be to further enhance the available size information, especially by adding stereoscopic information, while also reducing flatness cues. If, under such conditions, the main effect of environmental size ratio now strongly dominated, this would suggest that the "best estimate" interpretation is correct—with the estimate increasingly agreeing with the environmental size specified by the information in the displays as this information is enriched and conflicting flatness cues are removed. In the absence of this test, the best estimate account would seem more plausible for two reasons. First, even with incomplete and conflicting size information, the effect of environmental size ratio in the hallway was consistently higher than the effect of retinal (or flat) size ratio. Second, the proximal mode is unnatural and requires effort, and so it would seem an unlikely strategy to adopt in an experiment requiring speeded responses. However, even if the present results partially reflect a proximal mode, the large hallway environmental size ratio effects—and the significant "environmental" size ratio effect in the wall condition—still indicate a strong tendency to recover environmental size prior to the mental scaling and form matching.

Related Results

Related, complementary results have been obtained in a number of studies in which the stimuli, tasks, and theoretical motivations contrasted with those of the present study, in whole or in part.

On such area of research concerns the study of visual attention, using the visual search paradigm. There is accumulating evidence that some surface and scene information is accessible preattentively, at least in nascent form, as gauged by its availability to guide rapid search (Aks & Enns, 1992, 1996; Enns & Rensink, 1990, 1991; He & Nakayama, 1992; Kleffner & Ramachandran, 1992; Rensink & Enns, 1995; see also McLeod, Driver, & Crisp, 1988; Nakayama & Silverman, 1986, Ramachan-

dran, 1988). Several of these studies provide evidence for the strong conclusion that certain stimulus elements are not accessible preattentively if presented in contexts signaling, or helping to signal, surface characteristics—even when the same stimulus elements guide rapid search when presented in other contexts, or when shown in isolation (see He & Nakayama, 1992; Rensink & Enns, 1995; see also Enns & Rensink, 1991). Perhaps most relevant to the present study are recent reports that environmental size information affects visual search in a size discrimination task (Aks & Enns, 1996; Ramachandran, 1989, as cited in Enns & Rensink, 1991). As Aks and Enns (1996) point out, their subjects did not seem to have preattentive access to the retinal sizes of the outline drawings of cylinders used, at least when the outlined cylinders were shown against—and seen as attached to—receding texture planes.

Milliken and Jolicœur (1992) also obtained complementary results, using a variant of an old–new memory task to explore the representation of size as it is held in memory over several minutes. Other studies (Biederman & Cooper, 1992; Jolicœur, 1987; see also Cooper, Schacter, Ballesteros, & Moore, 1992) have shown that reaction time in an old–new memory task is lower when the test form is the same size as the earlier, studied form, as opposed to when the sizes differ. Milliken and Jolicœur prized apart retinal and environmental size by placing subjects at different distances from the computer screen (using two different retinal and environmental ratios, 1:1 and 2:2), and obtained evidence that it is "perceived size" that is represented in memory.

A further line of converging evidence has been provided by Burbeck (1987). Spatial-frequency discrimination is a widely used tool in the study of spatial vision, but Burbeck pointed out that since the sine wave gratings are shown at the same distance, retinal frequency (cycles/degree) and object frequency are confounded. Burbeck teased the two apart by presenting gratings on two monitors located at different distances. Perhaps most telling, subjects appeared unable to learn to make the discrimination on the basis of retinal frequency, suggesting a strong, natural tendency to represent and compare object frequencies instead.

Several studies of form and pattern perception have also yielded broadly complementary results. One lesson of the earlier size scaling experiments is that size is represented integrally with form during the matching process, and the present experiments reveal that it is environmental size that is represented, at least in part. This is in keeping with studies suggesting that the perception of form and pattern is fixed by the relation of parts or elements in nonretinal reference frames (Palmer, 1992; Rock & Brosgole, 1964; Rock & Linnett, 1993).

Finally, Uhlarik, Pringle, Jordan, and Misceo (1980) explored reaction times of verbal size estimates as the instructions, and the retinal and environmental sizes of objects, were varied. The stimuli used by Uhlarik et al. were photographs of single white blocks of different

sizes displayed at varying distances along a textured surface. Subjects were told to judge the size of the blocks in terms of a 10-unit standard block, shown separately. Under all instructions, Uhlarik et al. found that reaction time increased with increases in the environmental size ratio of the test object and the standard. Uhlarik et al. suggested that subjects proceed by forming a mental image of the standard and mentally scaling it to fit the test stimuli, somehow counting, or measuring off, the size of the test stimulus in terms of the standard (see also Hartley, 1977).⁴

One problem with using verbal ratings to study size perception is that performance is sensitive to slight differences in terms of the way the instructions are framed and/or understood. And Uhlarik et al. reported that 8 subjects were replaced because, after the experiment, "They verbalized judgmental modes other than the one instructed" (p. 63). Similar problems afflict methods that require the adjustment of a comparison object to indicate size judgments (Sedgwick, 1986). By contrast, the form matching task employed in Experiments 1 and 2 is simple, natural, and unambiguous. Although size is an *irrelevant* dimension, it nonetheless affects reaction time, and in a way that is sensitive to available size information. These considerations suggest that the form matching paradigm employed in the present study may provide a useful, alternative, indirect measure of size perception—perhaps especially if the "best estimate" account of subjects' performance on the task can be established as correct (see Aks & Enns, 1996, for a related application of the visual search paradigm; see also Aks & Enns, 1992).

Theoretical Implications

The present study was motivated in part by appeal to the theory of high-level vision advanced by Kosslyn et al. (1990). The fact that environmental size is, at least in part, recovered prior to mental scaling and form matching indicates that the theory needs to be revised. The results of Experiments 1 and 2 suggest that the environmental size of local, visually present surfaces is represented at a level roughly corresponding to the Kosslyn et al. "visual buffer." And the related work discussed in the preceding section suggests that a range of additional surface and scene information is also recovered early in visual processing.

It is, however, important to distinguish different experimental tasks, which may tap functionally distinct capacities. So, for example, in perceptually *classifying* an object, the aim is to attach a label to exemplars that share certain general structural characteristics, and size does not appear to affect performance (Biederman & Cooper, 1992; Cooper et al., 1992; see also Landau, Smith, & Jones, 1988). By contrast, it is important to extract and maintain information about environmental size in order to guide grasping and navigation. And the perceptual processes that help serve the latter functions may underlie performance on the same-different form matching task employed in Experiments 1 and 2, as well as the results obtained in the body of related work surveyed in the

previous section. It seems plausible that the size-discrimination visual search task (Aks & Enns, 1996; Ramachandran, 1989), the form matching task used in Experiments 1 and 2 of the present study, and the old-new memory task (Milliken & Jolicoeur, 1992) illuminate different stages of a connected process of size perception and recollection.

Finally, as noted, Kosslyn et al. (1990) proposed that size is coded in terms of visual angle at the level of the visual buffer because of the received view that the early visual system consists of retinotopic maps. However, to begin with, the evidence suggesting that there are *topographically* organized visual maps does not establish that these maps are *retinotopically* organized: The standard way of determining receptive fields and their spatial organization is by presenting isolated stimuli to immobilized animals (Hubel & Wiesel, 1962), but this confounds retinal and environmental spatial relations in much the same way that retinal and environmental size were confounded in the original size scaling experiments (see also, Pouget, Fisher, & Sejnowski, 1993). Further, though the functional significance is not entirely clear, recent experiments have revealed that the size and possibly "[visual field] location (Gilbert & Wiesel, 1992) of V1 receptive fields can change dynamically, within minutes, in response to retinal lesions (Gilbert & Wiesel, 1992) and to "artificial scotomas" (Pettet & Gilbert, 1992). It has, as well, long been known that stimulation outside the "classical receptive field" of visual cortex neurons can affect response to stimuli presented within the receptive field (see Gilbert & Wiesel, 1990; Knierim & Van Essen, 1992); Lamme (1995) and Zipser, Lamme, and Schiller (1996) have recently provided compelling evidence that, with a range of stimuli, such "contextually modified" responses of V1 neurons are linked to the perception of surfaces segregated from backgrounds. These results cast doubt on the view that the early visual system is best understood as consisting of stable, retinotopic "maps" of "feature detectors" with fixed response properties (at least exclusively—see Zipser et al., 1996). Rather, these experiments suggest that the early visual system contributes to the detection of basic surface properties.

Conclusions

The results of Experiments 1 and 2 suggest that prior to mental scaling and form matching, subjects make a best estimate of environmental size, given the available information. The paradigm developed shows promise as an indirect reaction time measure of size perception.

REFERENCES

- AKS, D. J., & ENNS, J. (1992). Visual search for direction of shading is influenced by apparent depth. *Perception & Psychophysics*, *52*, 63-74.
- AKS, D. J., & ENNS, J. (1996). Visual search for size is influenced by a background texture gradient. *Journal of Experimental Psychology: Human Perception & Performance*, *22*, 1467-1481.
- BESNER, D. (1983). Visual pattern recognition: Size preprocessing re-examined. *Quarterly Journal of Experimental Psychology*, *35A*, 209-216.

- BIEDERMAN, I., & COOPER, E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 121-133.
- BROOTA, K. D., & EPSTEIN, W. (1973). The time it takes to make veridical size and distance judgments. *Perception & Psychophysics*, **14**, 358-364.
- BUNDESEN, C., & LARSEN, A. (1975). Visual transformation of size. *Journal of Experimental Psychology: Human Perception & Performance*, **1**, 214-220.
- BUNDESEN, C., LARSEN, A., & FARRELL, J. E. (1981). Mental transformations of size and orientation. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 279-294). Hillsdale, NJ: Erlbaum.
- BURBECK, C. (1987). Locus of spatial-frequency discrimination. *Journal of the Optical Society of America A*, **4**, 1807-1813.
- CARLSEN, V. R. (1977). Instructions and perceptual constancy judgments. In W. Epstein (Ed.), *Stability and constancy in visual perception* (pp. 217-254). New York: Wiley.
- CAVE, K. R., & KOSSLYN, S. M. (1989). Varieties of size-specific visual selection. *Journal of Experimental Psychology: General*, **118**, 148-164.
- CAVE, K. R., PINKER, S., GIORGI, L., THOMAS, C. E., HELLER, L. M., WOLFE, J., & LIN, H. (1994). The representation of location in visual images. *Cognitive Psychology*, **26**, 1-32.
- COOPER, L. A., SCHACTER, D. L., BALLESTEROS, S., & MOORE, C. (1992). Priming and recognition of transformed three-dimensional objects: Effects of size and reflection. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 43-57.
- CORCORAN, D. W. J., & BESNER, D. (1975). Application of the Posner technique to the study of size and brightness irrelevancies in letter pairs. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance V* (pp. 613-630). New York: Academic Press.
- CUTTING, J. E., & VISHON, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein & S. Rogers (Eds.), *Perception of space and motion: Handbook of perception and cognition* (2nd ed., pp. 69-117). San Diego: Academic Press.
- ELLIS, R., ALLPORT, D. A., HUMPHREYS, G. W., & COLLIS, J. (1989). Varieties of object constancy. *Quarterly Journal of Experimental Psychology*, **41A**, 775-796.
- ENNS, J. T., & RENSINK, R. A. (1990). Sensitivity to three-dimensional orientation in visual search. *Psychological Science*, **1**, 323-326.
- ENNS, J. T., & RENSINK, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review*, **98**, 335-351.
- EPSTEIN, W. (1963). Attitudes of adjustment and the size-distance invariance hypothesis. *Journal of Experimental Psychology*, **66**, 78-83.
- EPSTEIN, W., & BROOTA, K. D. (1975). Attitude of judgment and reaction time in estimation of size at a distance. *Perception & Psychophysics*, **18**, 201-204.
- FARAH, M. (1990). The neural basis of mental imagery. *Trends in Neurosciences*, **12**, 461-470.
- FARRELL, B. (1985). "Same-different" judgments: A review of current controversies in perceptual comparisons. *Psychological Bulletin*, **98**, 419-456.
- FINKE, R. A. (1989). *Principles of mental imagery*. Cambridge, MA: MIT Press.
- FOLEY, J. D., VAN DAM, A., FEINER, S. K., & HUGHES, J. F. (1990). *Computer graphics: Principles and practice* (2nd ed.). New York: Addison-Wesley.
- FOLK, M. D., & LUCE, R. D. (1987). Effects of stimulus complexity on mental rotation of polygons. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 395-404.
- GIBSON, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.
- GILBERT, C. D., & WIESEL, T. (1990). The influence of contextual stimuli on the orientation selectivity of cells in the primary visual cortex of the cat. *Vision Research*, **30**, 1689-1701.
- GILBERT, C. D., & WIESEL, T. (1992). Receptive field dynamics in adult primary visual cortex. *Nature*, **356**, 150-152.
- GILLAM, B. (1995). The perception of spatial layout from static optical information. In W. Epstein & S. Rogers (Eds.), *Perception of space and motion: Handbook of perception and cognition* (2nd ed.). San Diego: Academic Press.
- GOGEL, W. C. (1990). A theory of phenomenal geometry and its applications. *Perception & Psychophysics*, **48**, 105-123.
- GREENHOUSE, S. W., & GEISSER, S. (1959). On methods in analysis of profile data. *Psychometrika*, **24**, 95-112.
- HARTLEY, A. A. (1977). Mental measurement in the magnitude estimation of length. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 622-628.
- HE, Z. J., & NAKAYAMA, K. (1992). Surface versus features in visual search. *Nature*, **359**, 231-233.
- HOCHBERG, J. (1971). Perception II. Space and movement. In J. W. Kling & L. A. Riggs (Eds.), *Woodworth and Schlosberg's Experimental psychology* (pp. 475-550). New York: Holt, Rinehart & Winston.
- HOWARD, J. H., JR., & KERST, S. M. (1978). Directional effects of size change on the comparison of visual shapes. *American Journal of Psychology*, **91**, 491-499.
- HUBEL, D. H., & WIESEL, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, **160**, 106-154.
- JOLICŒUR, P. (1987). A size-congruency effect in memory for visual shape. *Memory & Cognition*, **15**, 531-543.
- JOLICŒUR, P., & BESNER, D. (1987). Additivity and interaction between size ratio and response category in the comparison of size-discrepant shapes. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 478-487.
- KERSTEN, D., MAMMASSIAN, P., & KNILL, D. (1997). Moving cast shadows induce apparent motion in depth. *Perception*, **26**, 171-192.
- KLEFFNER, D. A., & RAMACHANDRAN, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics*, **52**, 18-36.
- KNIERIM, J. J., & VAN ESSEN, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology*, **67**, 961-980.
- KOSSLYN, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.
- KOSSLYN, S. M. (1987). Seeing and imaging in the cerebral hemispheres: A computational approach. *Psychological Review*, **94**, 148-175.
- KOSSLYN, S. M., FLYNN, R. A., AMSTERDAM, J. B., & WANG, G. (1990). Components of high-level vision: A cognitive neuroscience analysis and accounts of neurological syndromes. *Cognition*, **34**, 203-277.
- LAMME, V. A. F. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience*, **15**, 1605-1615.
- LANDAU, B., SMITH, L., & JONES, S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, **3**, 299-321.
- LARSEN, A. (1985). Pattern matching: Effects of size ratio, angular difference in orientation, and familiarity. *Perception & Psychophysics*, **38**, 63-68.
- LARSEN, A., & BUNDESEN, C. (1978). Size scaling in human pattern recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 1-20.
- MARK, L. S. (1987). Eyeheight-scaled information about affordances: A study of sitting and stair climbing. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 361-370.
- MARR, D. (1982). *Vision*. San Francisco: Freeman.
- McKEE, S. P., & SMALLMAN, H. S. (1998). Size and speed constancy. In V. Walsh & J. Kulikowski (Eds.), *Perceptual constancy* (pp. 373-408). Cambridge: Cambridge University Press.
- McKEE, S. P., & WELCH, L. (1992). The precision of size constancy. *Vision Research*, **32**, 1447-1460.
- MCLEOD, P., DRIVER, J., & CRISP, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature*, **320**, 264-265.
- MILLIKEN, B., & JOLICŒUR, P. (1992). Size effects in visual recognition memory are determined by perceived size. *Memory & Cognition*, **20**, 83-95.
- NAKAYAMA, K., & SILVERMAN, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, **320**, 264-265.
- PALMER, S. E. (1992). Common region: A new principle of perceptual grouping. *Cognitive Psychology*, **24**, 436-447.
- PETTET, M. W., & GILBERT, C. D. (1992). Dynamic changes in receptive-field size in cat primary visual cortex. *Proceedings of the National Academy of Sciences*, **89**, 8366-8370.
- POSNER, M. I., & MITCHELL, R. F. (1967). Chronometric analysis of classification. *Psychological Review*, **74**, 392-409.

- POUGET, A., FISHER, S. A., & SEJNOWSKI, T. J. (1993). Egocentric spatial representation in early vision. *Journal of Cognitive Neuroscience*, **5**, 150-161.
- PURDY, W. C. (1958). *The hypothesis of psychophysical correspondence in space perception*. Unpublished doctoral dissertation, Cornell University.
- RAMACHANDRAN, V. S. (1988). Perceiving shape from shading. *Scientific American*, **259**, 76-83.
- RAMACHANDRAN, V. S. (1989, November). *Is perceived size computed before or after visual search?* Paper presented at the meeting of the Psychonomic Society, Atlanta.
- RENSINK, R. A., & ENNS, J. T. (1995). Preemption effects in visual search: Evidence for low-level grouping. *Psychological Review*, **102**, 101-130.
- ROCK, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.
- ROCK, I., & BROSGOLE, L. (1964). Grouping based on phenomenal proximity. *Journal of Experimental Psychology*, **67**, 531-538.
- ROCK, I., & LINNETT, C. M. (1993). Is a perceived shape based on its retinal image? *Perception*, **22**, 61-76.
- SEDGWICK, H. A. (1980). The geometry of spatial layout in pictorial representation. In M. A. Hagen (Ed.), *The perception of pictures. Vol. 1: Alberti's window* (pp. 33-90). New York: Academic Press.
- SEDGWICK, H. A. (1986). Space perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance. Vol. 1: Sensory processes and perception* (pp. 21.1-21.57). New York: Wiley.
- SEKULAR, R., & NASH, D. (1972). Speed of size-scaling in human vision. *Psychonomic Science*, **27**, 93-94.
- SHEPARD, R. N., & METZLER, J. (1971). Mental rotation of three-dimensional objects. *Science*, **171**, 701-703.
- STEVENS, K. A. (1979). Representing and analyzing surface orientation. In P. H. Winston & R. H. Brown (Eds.), *Artificial intelligence: An MIT perspective* (Vol. 1, pp. 103-129). Cambridge, MA: MIT Press.
- STEVENS, K. A. (1981). The information content of texture gradients. *Biological Cybernetics*, **42**, 95-105.
- THOULESS, R. H. (1931). Phenomenal regression to the real object: I. *British Journal of Psychology*, **21**, 339-359.
- UHLARIK, J., PRINGLE, R., JORDAN, K., & MISCEO, G. (1980). Size scaling in two-dimensional pictorial arrays. *Perception & Psychophysics*, **27**, 60-70.
- WALLACH, H., & O'LEARY, A. (1982). Slope of regard as a distance cue. *Perception & Psychophysics*, **31**, 145-148.
- WARREN, W. H., & WHANG, S. (1987). Visual guidance of walking through apertures: Body-scaled information for affordances. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 371-383.
- WRAGA, M. (1999a). The role of eye height in perceiving affordances and object dimensions. *Perception & Psychophysics*, **61**, 490-507.
- WRAGA, M. (1999b). Using of eye height in different postures to scale the heights of objects. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 518-530.
- YONAS, A. (1978). Development of sensitivity to information provided by cast shadows in pictures. *Perception*, **7**, 333-341.
- ZIPSER, K., LAMME, V. A. F., & SCHILLER, P. H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience*, **16**, 7376-7389.

NOTES

1. Cave and Kosslyn (1989) showed that other functions also fit their data and Experiment 2 of Larsen and Bundesen (1978), but they, too, found a close to linear dependence of reaction time on size ratio.

2. Polygon anti-aliasing was used only on the random forms, and not the background scene (i.e., the hallway or wall). Ordinarily, if VGX polygon anti-aliasing is used to render an isolated form, the internal edges of the triangles making up the form are visible. This problem was averted by layering a slightly smaller, unaliased form of the same shape over the anti-aliased form. Although on close inspection several lighter pixels were sometimes visible at the vertices of the random forms, no subjects reported noticing this, and the immediate perception was of a uniformly black form, with sharp, crisp edges.

3. Due to experimenter error, one session of 1 subject was run out of order. The first four sessions for this subject were shadows, hall, hall, shadows. For purposes of counterbalancing, this subject was counted as running the shadows condition first, and the hall condition second, since this was the order of the initial exposure to the two conditions. Although the possible presence of order effects was not examined in detail, the environmental size slope in the hall condition was actually slightly lower for those 6 subjects who ran in this condition first, than for those 6 subjects who ran in the wall condition first—suggesting that, for these stimuli, there was little if any effect of order.

4. Epstein and Broota (1975; see also Broota & Epstein, 1973) also examined reaction times of judgments of the size of rectangles in inches (subjects were shown a standard foot rule at the beginning of each session). Although no effect of environmental size was reported, the graph of the results of Experiment 1 of Epstein and Broota is consistent with such an effect, and it does not appear that the data was analyzed in a way that would decide this issue. Epstein and Broota did find an effect of distance on reaction time, which conflicts with the failure to find an effect of distance by Uhlarik et al. (1980); however, there were a number of methodological differences between the two studies, some discussed by Uhlarik et al.

(Manuscript received February 7, 2001;
revision accepted for publication June 4, 2001.)