On the perception of shape from shading

DOROTHY A. KLEFFNER and V. S. RAMACHANDRAN
University of California, San Diego, La Jolla, California

The extraction of three-dimensional shape from shading is one of the most perceptually compelling, yet poorly understood, aspects of visual perception. In this paper, we report several new experiments on the manner in which the perception of shape from shading interacts with other visual processes such as perceptual grouping, preattentive search ("pop-out"), and motion perception. Our specific findings are as follows: (1) The extraction of shape from shading information incorporates at least two "assumptions" or constraints—first, that there is a single light source illuminating the whole scene, and second, that the light is shining from "above" in relation to retinal coordinates. (2) Tokens defined by shading can serve as a basis for perceptual grouping and segregation. (3) Reaction time for detecting a single convex shape does not increase with the number of items in the display. This "pop-out" effect must be based on shading rather than on differences in luminance polarity, since neither left-right differences nor step changes in luminance resulted in pop-out. (4) When the subjects were experienced, there were no search asymmetries for convex as opposed to concave tokens, but when the subjects were naive, cavities were much easier to detect than convex shapes. (5) The extraction of shape from shading can also provide an input to motion perception. And finally, (6) the assumption of "overhead illumination" that leads to perceptual grouping depends primarily on retinal rather than on "phenomenal" or gravitational coordinates. Taken collectively, these findings imply that the extraction of shape from shading is an "early" visual process that occurs prior to perceptual grouping, motion perception, and vestibular (as well as "cognitive") correction for head tilt. Hence, there may be neural elements very early in visual processing that are specialized for the extraction of shape from shading.

We use three-dimensional (3-D) depth perception to find our way around the world and to manipulate objects that we encounter. Although the retinal image is two-dimensional, somehow the brain is able to use the information from this image to yield an experience of solidity and depth.

Of the numerous mechanisms used by the visual system to recover the third dimension, the ability to use shading is probably phylogenetically one of the most primitive. One reason for believing this is that in the natural world, animals have often evolved the principle of countershading to conceal their shapes from predators; they have pale bellies that serve to neutralize the effects of the sun shining from above (Thayer, 1909). The prevalence of countershading in a variety of animals (including fishes) suggests that shading must be a very important source of information about 3-D shapes.

Although artists have long recognized the importance of shading, there have been few studies of how the human visual system actually extracts and uses this information. Since the time when Leonardo da Vinci first

thought about this problem, there have been only a small handful of systematic psychological studies on it (Berbaum, Bever, & Chung, 1983; Brewster, 1847; Howard, 1983; Ramachandran, 1988a, 1988b; Rittenhouse, 1786; Todd & Mingolla, 1983).

We began our investigations by creating a set of simple computer-generated displays (Figure 1). The impression of depth perceived in these displays is based exclusively on subtle variations in shading that we made sure were devoid of any complex objects and patterns. Our purpose, of course, was to isolate the brain mechanisms that process shading information from other mechanisms that may also contribute to depth perception in real-life visual processing. So the displays are intended to serve the same role in the study of shape from shading that Julesz's stereograms (Julesz, 1971) do in the study of stereopsis.

We have recently used these computer-generated displays to discover a simple set of "rules" or constraints that the visual system uses in the interpretation of 3-D shape from shading (Ramachandran 1988a, 1988b). For example, Figure 1 depicts a set of objects that conveys a strong impression of depth. The sign of perceived depth, however, is ambiguous, since the visual system has no way of knowing where the light source is. Consequently, the display can be perceived as consisting of either convex objects illuminated from the right or concave objects lit from the left ("eggs" or "egg-crate"). The reader can generate a depth inversion as though mentally "shifting" the light source.
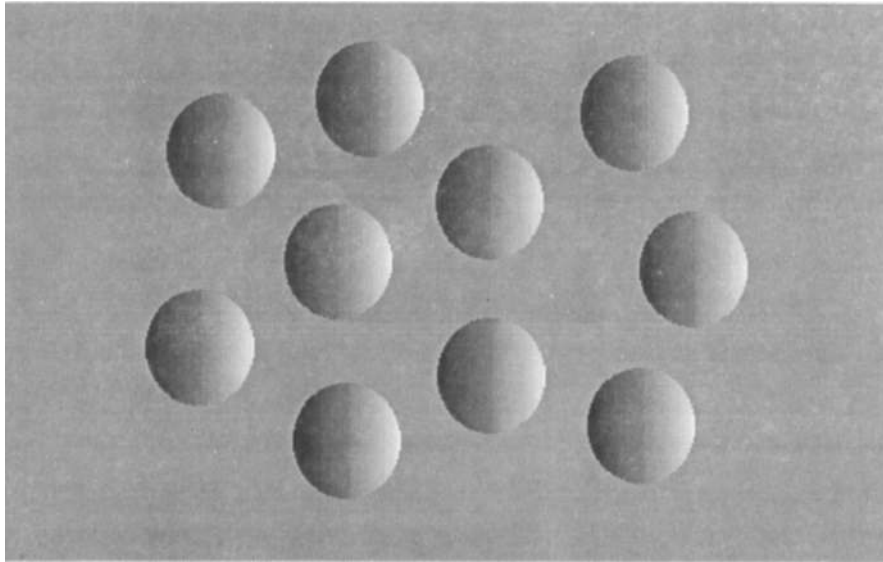
18

Figure 1. These computer-generated displays convey an impression of depth based exclusively on subtle variations in luminance. The sign of perceived depth is ambiguous. Each object can be perceived as either convex and lit from the right or concave and lit from the left, but all of the objects tend to be viewed with the same sign of perceived depth.

Interestingly, when a depth inversion occurs, it tends to occur simultaneously for all objects in the display. Is this propensity for seeing all objects in the display as being simultaneously convex (or concave) based on a tendency to assign identical depth values to all of them, or is it based on the tacit assumption that there is only *one* light source in the image? To find out, we used a mixture of objects that were mirror images of each other (Figure 2). In this display, when the top row of objects was seen as convex, the bottom row was always perceived as concave, and vice versa. It was in fact impossible to see all the objects as being simultaneously convex or concave. This observation suggests that when interpreting shape from shading, the visual system incorporates the tacit assumption that there is only one light source illuminating the entire visual image (or a large portion of it; Ramachandran, 1988b). Hence the derivation of shape from shading cannot be a strictly local operation; it must involve "global" assumptions about light sources.

Note that, as in Figure 2, a row can be seen as either convex or concave if the other row is excluded. When both rows are viewed simultaneously, however, seeing one row as convex forces the other row to be perceived as concave. Some powerful inhibitory mechanisms must be involved in the generation of these effects. The single-light-source assumption is, of course, implicit in many
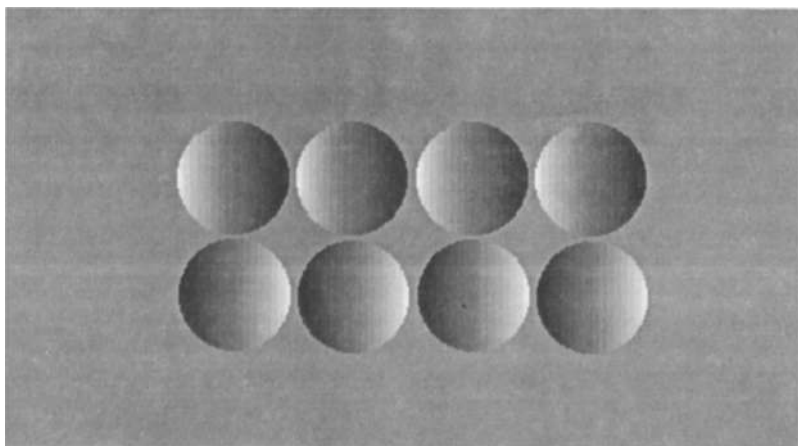

Figure 2. The single-light-source assumption, demonstrated through the use of a mixture of shaded objects that are mirror images of each other: Objects in one row can be seen as either convex or concave if the other row is excluded; but when both rows are viewed simultaneously, seeing one row as convex forces the other row to be perceived as concave.
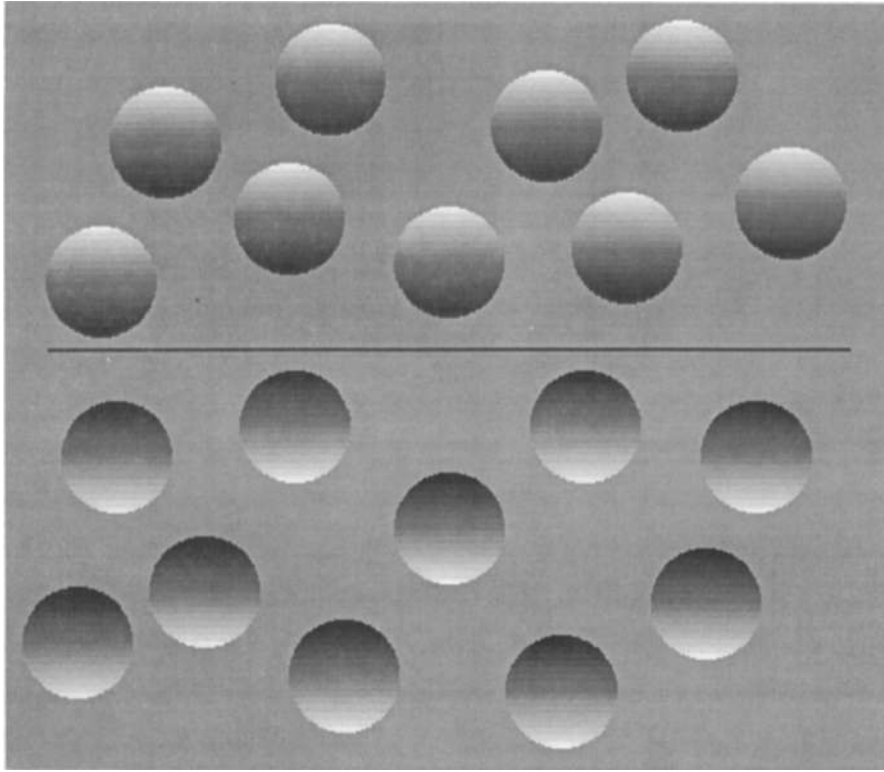
Figure 3. This computer-generated photograph demonstrates that the visual system has a built-in "assumption" that the light source is shining from above. Note that the depth in these displays is conveyed exclusively through shading, with no other depth cues present. The shaded objects in the top panel are usually seen as convex, whereas those in the bottom panel are usually seen as concave. Note, however, that the illusion (i.e., the difference between convex and concave) is not as pronounced as it is in Figure 5A, in which the objects are intermixed.

artificial intelligence models, but Figure 2, as far as we know, is the first clear-cut demonstration that such a rule actually exists in human vision for the extraction of shape from shading. Bergstrom (1987) has pointed out that such a rule may also be involved in the computation of surface lightness.

In addition to the single-light-source constraint described above, there appears also to be a built-in assumption that the light is shining from above, a principle first suggested by Sir David Brewster (1847). This would explain why, in Figure 3, objects in the top panel are usually seen as convex, whereas those in the bottom panel are often perceived as "holes" or "cavities." The sign of depth can be readily reversed by simply turning the figure upside down. The effect is weak, however, since either panel can be seen as convex if the other is excluded from view to eliminate the single-light-source constraint. On the other hand, if a mixture of such objects is presented, it is almost impossible to reverse any of them because of the combined effect of two constraints—the single-light-source constraint and the "top"-light-source constraint (Figure 5A).

Next, we wondered what would happen to the interpretation of shape from shading if one were to give the visual

system *conflicting* information about the light source's location. To explore this, we created the display shown in Figure 4. The central disks are identical in A and B, with a vertical gradient. The surround in A has a conflicting horizontal gradient, which could not occur with a single light source illuminating the display. The figure was shown to 48 naive subjects, who were asked to examine the two panels (A and B) carefully and compare the two central disks. Their task was to judge which of the two central figures (A or B) appeared more convex. The results were clear-cut; the central disk in panel B almost always appeared to be more convex than did the central disk in panel A (72 out of 96 trials). In fact, many subjects spontaneously reported that the disk in panel A almost appeared flat. We may conclude, therefore, that the magnitude of depth perceived from shading is enhanced considerably if objects in the surround have the opposite polarity, a spatial contrast effect that is vaguely reminiscent of the center-surround effects that have been reported for other stimulus dimensions such as motion (Nakayama & Loomis, 1974) and color (Land, 1983; Livingstone & Hubel, 1987). Another way of saying this would be that the perception of shape from shading is enhanced considerably if the information in the scene is com-
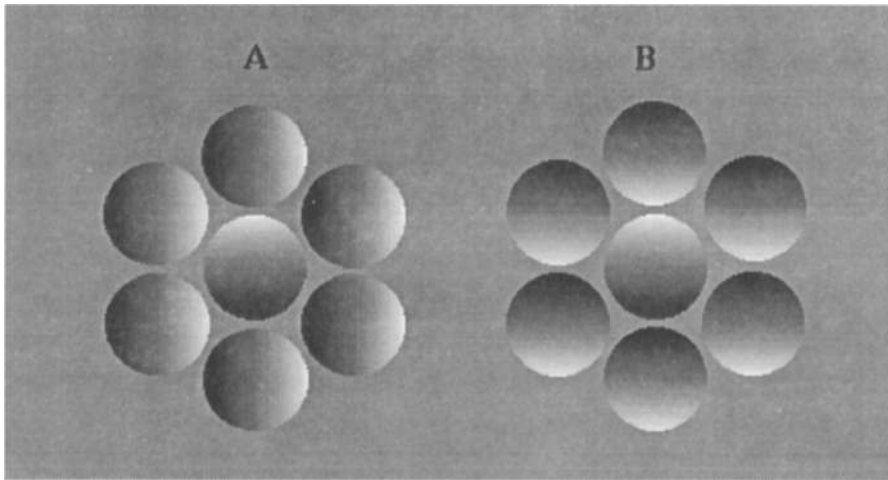
Figure 4. This display demonstrates "center-surround" interactions in the perception of shape from shading. The central disk in panel A is usually seen as less convex than the identical one in panel B. These effects are usually much more pronounced on the CRT than they are in the printed versions shown here. This effect demonstrates that the magnitude of perceived depth is also influenced by the single-light-source constraint (after Ramachandran, 1989b).

patible with a single light source. When the information from the majority of objects (e.g., panel A) suggests that the light source is on the left (or right), the shading on the central object is perceived as a variation in reflectance rather than depth (Ramachandran, 1989b).

What if the location of the light source was revealed by some obvious means? This question was first raised by Berbaum et al. (1983). They asked subjects to view a muffin pan illuminated from below while holding a hand nearby to cast a shadow—thereby revealing the light source. Berbaum et al. found that many subjects now reported a reversal of relief. Oddly enough, we did not find this to be true for our computer-generated displays. A hollow mask lit from above looks like a "normal" (convex) face lit from below. But if the eggs and cavities in Figure 5A are placed right next to it, their depth does not reverse (Ramachandran, 1988a), in spite of the fact that the face now "reveals" the light to be coming from below. Yet we found that if the eggs and cavities are directly pasted on the face with their outlines blurred in order to "blend" them into the face, then their depth does indeed reverse (i.e., the eggs become cavities, and vice versa). We may conclude, therefore, that the knowledge about the new light source location, revealed by the face, does not generalize to apply to other items in the display unless these items are seen as belonging to the face—that is, as being parts of the same object. Or, to put it differently, the single-light-source rule is adhered to more rigidly for different parts of an object than it is for different objects in a scene.

Note that it is also possible to group all the convex shapes in Figure 5A together mentally to form a cluster that is clearly segregated from the background of concave shapes. This result is surprising, for it is usually assumed that only certain elementary stimulus features such

as orientation, color, and "terminators" can be grouped together in this way (Beck, 1966; Julesz, 1971; Treisman, 1985, 1986). Figure 5A shows that even 3-D shapes that are conveyed by shading can provide tokens for perceptual grouping and segregation (Ramachandran, 1988a, 1988b). To make sure that the effect was not due to some more elementary image feature (such as luminance polarity), we produced a control stimulus (Figure 5B), in which the targets were similar to those in Figure 5A in terms of luminance polarity but did not convey any depth. In this display, it is difficult to segregate the tokens on the basis of differences in polarity, suggesting that the effects observed in Figure 5A must be based on 3-D shapes. Segregation is also much more pronounced for top-down differences in illumination than for left-right differences. For instance, if Figure 5A is rotated by 90°, the degree of segregation is also reduced correspondingly. This further supports the view that the effect depends on the 3-D shapes of the tokens rather than on luminance polarity (Ramachandran, 1988a, 1988b).

Our purpose in the rest of this communication is to describe some formal experiments that we carried out to confirm and extend our earlier observations (Kleffner & Ramachandran, 1989; Ramachandran, 1988a, 1988b).

Our preliminary observations, described in Figure 5A, suggested that shape from shading can serve as an elementary feature for perceptual grouping; but would the same results also hold for effortless preattentive search or "pop-out"? Consider the case of a single egg displayed against a background of several cavities. The extent to which reaction times vary with the number of items in a display is often used as a criterion to decide whether a particular visual feature is detected "preattentively" or not. If subjects do not have to search for the target—that is, if they can spot it without inspecting every item on display—
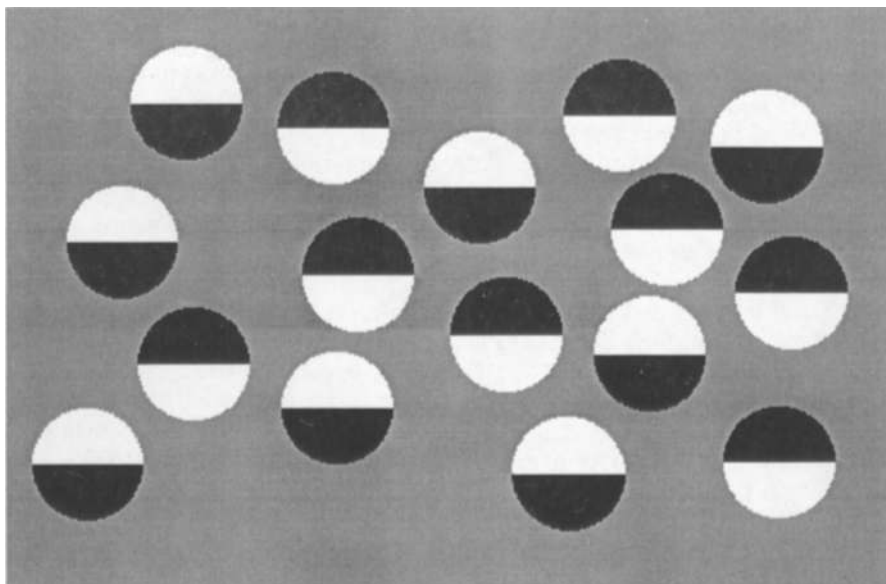
Ⓐ

Ⓑ
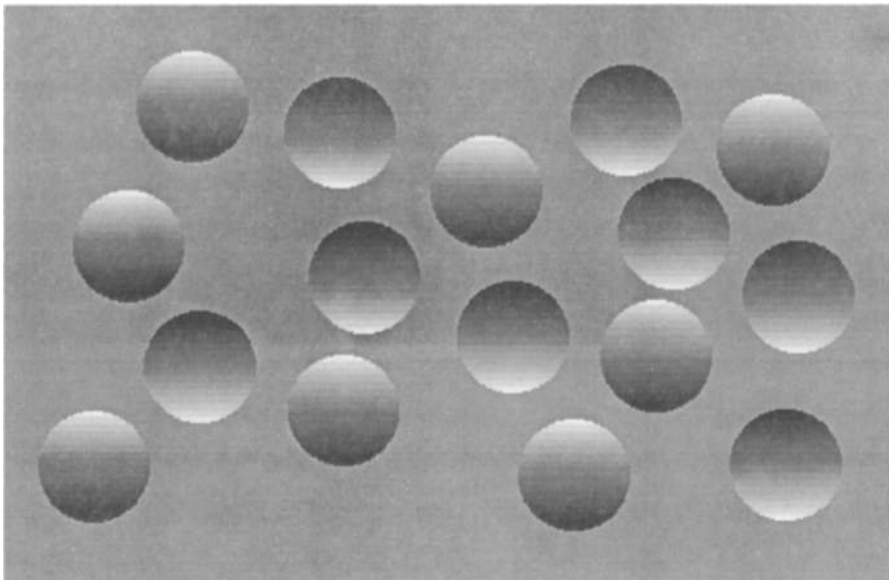
Figure 5. (A) This figure contains a random mixture of shaded objects that have opposite luminance polarities. The ones that are light on top are usually perceived as spheres that can be mentally grouped together and segregated from the background of concave objects. Hence we may conclude that three-dimensional shapes defined by shading can provide tokens for perceptual grouping and segregation. If the figure is rotated 90°, segregation becomes much more difficult. (B) Tokens in this control display have the same luminance polarity as the shaded images do, but they do not convey depth information. Segregation of the tokens is difficult to achieve.

then the feature in question is, by definition, "elementary." The reaction time for spotting such a target will not increase linearly with the number of distractors (Treisman, 1985, 1986). We decided to use this criterion to find out whether or not an egg would appear to pop out against a background of cavities. Subjects were simply asked to report the presence or absence of a single egg against a background consisting of a varying number of distractors (cavities).

EXPERIMENT 1
Visual Search for 3-D Shape from Shading

Method

Subjects. Five subjects participated: the 2 authors, 1 other researcher in the lab, and 2 undergraduate research assistants.

Display. This display, as well as subsequent ones described in this paper, were all generated on a CRT driven by an Amiga microcomputer. The targets and distractors are illustrated in Figure 6. Targets and distractor items subtended 1.0° of visual angle and were placed in random positions without overlap, within a display area 6.1° high × 6.6° wide. On each trial, 1, 6, or 12 items were displayed, with half the trials containing one target and the remaining trials containing no targets. Targets and distractors were constructed from 16 luminance levels ranging from .057 to 136.1 cd/m² and presented on a background of 14.6 cd/m².

Procedure. The subjects were seated .75 m from the screen in a dark room. Each trial began with a dark screen (.057 cd/m²) for 0.8 sec, followed by the presentation of a fixation point on a gray background (0.76 cd/m²) for 1.8 sec. The experimental stimulus was then displayed. Two keys on the keyboard were used by the subjects to indicate whether the target was present or absent in the display, and the subjects' reaction times were recorded. A response from the subject ended the trial, and the screen was once again blacked out. Subjects were given feedback after each trial, consisting of a "+" or "−" on a blank screen, which indicated whether or not the response was correct. This also served as the fixation point for the next trial.

Each block of the experiment consisted of 48 trials presented in random order, 8 trials from each of 6 conditions (1, 6, or 12 total items, with the target item either present or absent). The subjects completed four experimental blocks for each target-distractor set. Prior to the collection of data for each condition, the subjects practiced the experiment with the test stimulus until they felt comfortable (this was done for at least one block, but for less than four blocks).

Results

The major findings of this study were that subjects' ability to detect targets shaded vertically was significantly different from their ability to detect either horizontal shading or a step change in luminance. These results are shown in Figure 7. In the first display, with shading from top to bottom, reaction times were not dependent on the number of items in the display (Kleffner & Ramachandran, 1989). The slopes from this graph indicate an average reaction time of 4 msec per item when the target was present and 5 msec per item when the target was absent. But for the second display, shaded from left to right, reaction times did increase with the number of items in the display, to 22 msec per item for the target present condition and 50 msec per item for the target absent condition. This difference in slopes suggests that a "serial search" strategy was being used. The third display, a step change in luminance, gave mixed results. For most subjects, the reaction times varied with the number of distractor items in the display, but there was substantial variability between subjects. The average reaction time was 8 msec per item for the target present condition and 18 msec per item for the target absent condition.

A statistical comparison was made of the resulting slopes (reaction time vs. number of items in the display) from the graphs in Figure 7. A two-way analysis of variance (ANOVA) with repeated measures was performed, with the line slopes from the graphs as the dependent vari-
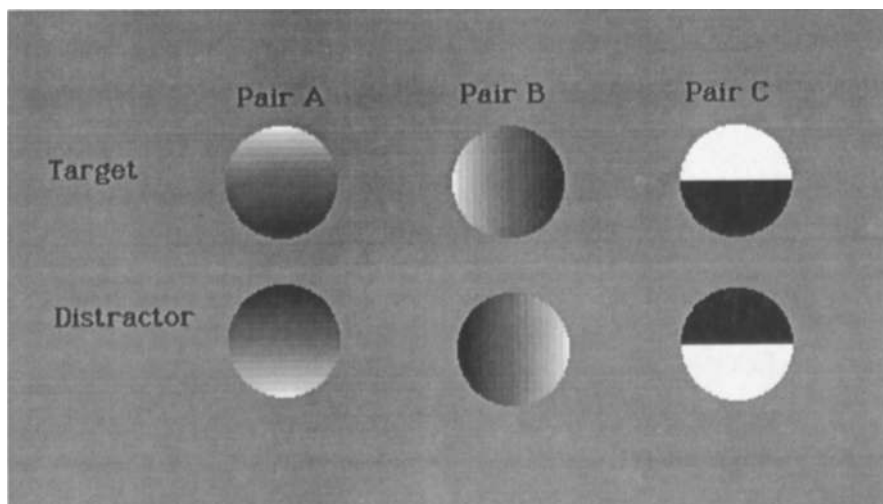

Figure 6. Examples of the target-distractor sets used in Experiment 1. (A) An object shaded top to bottom had to be detected against a field of distractors shaded from bottom to top. (B) An object shaded from left to right had to be detected against distractors that were shaded right to left. (C) A step change in luminance in the vertical direction had to be detected against a background of distractors that had the opposite polarity.

Ⓐ

Visual Search Task
Vertical Luminance Gradients

Ⓑ

Visual Search Task
Horizontal Luminance Gradients

Ⓒ

Visual Search Task
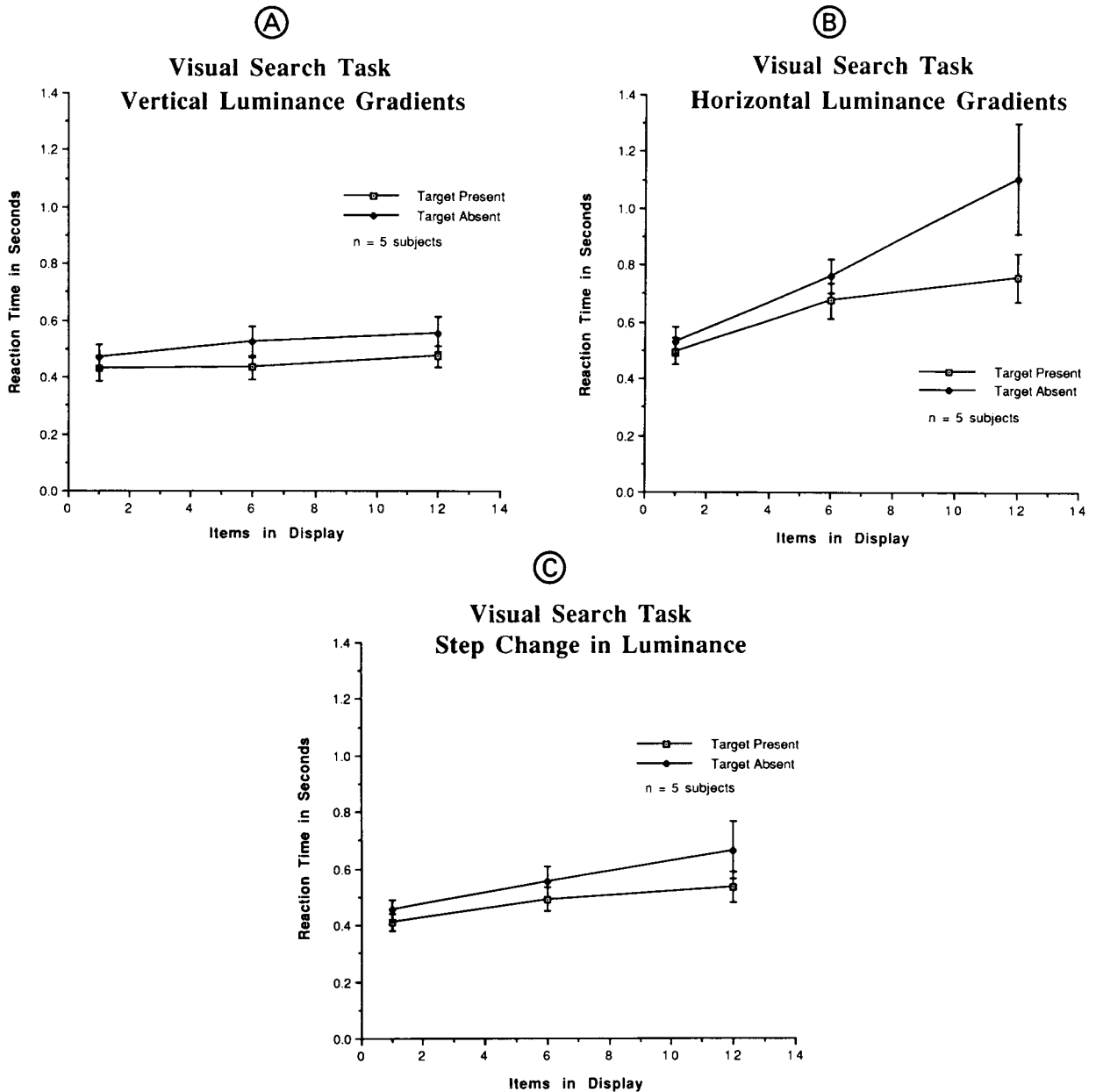## Step Change in Luminance

Figure 7. Results obtained from the visual search task, in which 5 experienced subjects participated. For vertical shading (A), the reaction time is unaffected by the number of distractors in the display. For horizontal shading (B), however, subjects' reaction times increased monotonically with the number of items in the display. When the stimulus was a step change in luminance (C), reaction time generally increased with the number of items in the display, but there was considerable variability between subjects.

able. The main effect for target type was significant at the .01 level [$F(2,16) = 8.129, p < .0038$], indicating that subjects' performance was significantly different in the three experimental conditions. (The second factor in the ANOVA, whether the target was present or absent in each trial, was included in the analysis to account for variance. This factor, and the interaction between the factors, was not significant here or in the following three comparisons.) ANOVAs were also used to make a direct comparison between pairs of experimental conditions. In a comparison of top/bottom shading with left/right shading, the main effect for target type was significant at the

.05 level [$F(1,8) = 10.886, p < .011$]. Top/bottom shading against a step change in luminance also produced a significant main effect for target type at the .05 level [$F(1,8) = 9.058, p < .0168$]. The difference between left/right shading and a step change in luminance, on the other hand, was not significant.

Discussion

These results suggest that the extraction of shape from shading can provide a basis for effortless or "preattentive" visual search, since reaction times do not increase with the number of distractors. The fact that such pop-

out is seen only for top–bottom differences in shading, and not for left–right differences, has two important implications. First, it implies that the effect must be based on the extraction of 3-D shape from shading, not just from differences in luminance polarity. Second, the process must incorporate the assumption that the light is shining from above. Hence certain "scene-based" image characteristics—such as the assumed location of light sources—can influence visual search (Ramachandran, 1988a, 1988b), a point that has also been elegantly demonstrated in the recent experiments of Enns and Rensink (1990). One anomalous finding is that the target defined by a step change in luminance also seemed to pop out more than one would expect from a casual inspection of Figure 5B. The reason for this might be that even though no depth is visible in this display, the mere presence of a vertical luminance gradient (with white on top) is sufficient to stimulate whatever neural detectors are involved in signaling convexity. The neurons may be excited suboptimally so that although the signal is strong enough to be detected in a search task, it is not strong enough to actually evoke a compelling sense of depth.

EXPERIMENT 2
Asymmetries in Visual Search

Treisman and Gormican (1988) noted that it is easier to detect a "closed" circle against a background of Cs (open circles) than it is to detect a C against a background of Os. They point out that such search asymmetries exist for a wide range of other types of visual features as well. Prompted by suggestions from A. Treisman and J. T. Enns, we decided to look for search asymmetries in the detection of 3-D shape from shading. In some preliminary experiments with experienced subjects, we found no evidence for an asymmetry, but we decided to repeat the experiments on naive subjects.

Method
Subjects. Six subjects from the undergraduate subject pool at the University of California, San Diego, participated in each of the conditions of the experiment (18 subjects total).

Display. The displays were identical to those in Experiment 1, with the exception that the target and distractor items were distinguished by top versus bottom shading, bottom versus top shading, and left versus right shading. These are shown in Figure 8.

Procedure. The procedure was identical to that used in Experiment 1, except that each set of 6 subjects participated in only one of the conditions (top vs. bottom shading, bottom vs. top shading, and left vs. right shading). Comparisons were therefore made across subjects rather than within subjects.

Results
The results (see Figure 9) showed a striking asymmetry. Surprisingly, it was much easier to detect a cavity against a background of eggs than vice versa.[1] For detecting an egg, reaction times increased with the number of items in the display, suggesting serial search. The average reaction time was 26 msec per item when the target was present and 50 msec per item when the target was absent. For detecting a cavity, however, reaction times did not increase with the number of items in the display. Average reaction times were 5 msec per item for both target present and target absent conditions. For the third display, which consisted of left to right shading, reaction times were again dependent on the number of items in the display—25 msec per item when the target was present and 60 msec per item when the target was absent.

A comparison was made between the resulting graphs (plotting reaction time vs. number of items in the display). A two-way ANOVA without repeated measures was performed, with the line slopes from the graphs as the dependent variable. The main effect for target type was significant at the .0001 level [$F(2,30) = 15.314, p < .0001$], indicating that the subjects' performance was significantly different in the three experimental conditions. The second factor, target presence/absence, was included in the
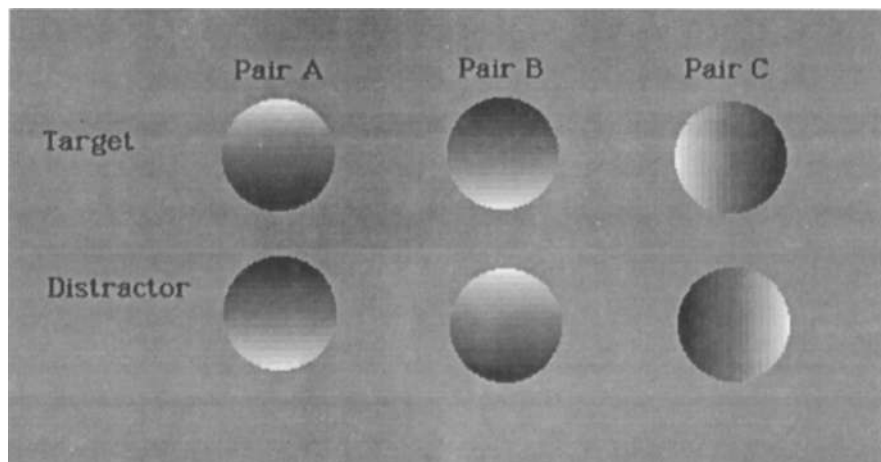

Figure 8. Examples of the target–distractor sets used in Experiment 2: (A) An object shaded top to bottom had to be detected against a field of distractors shaded from bottom to top. (B) An object shaded from bottom to top had to be detected against distractors that were shaded top to bottom. (C) An object shaded from left to right had to be detected against distractors that were shaded right to left.
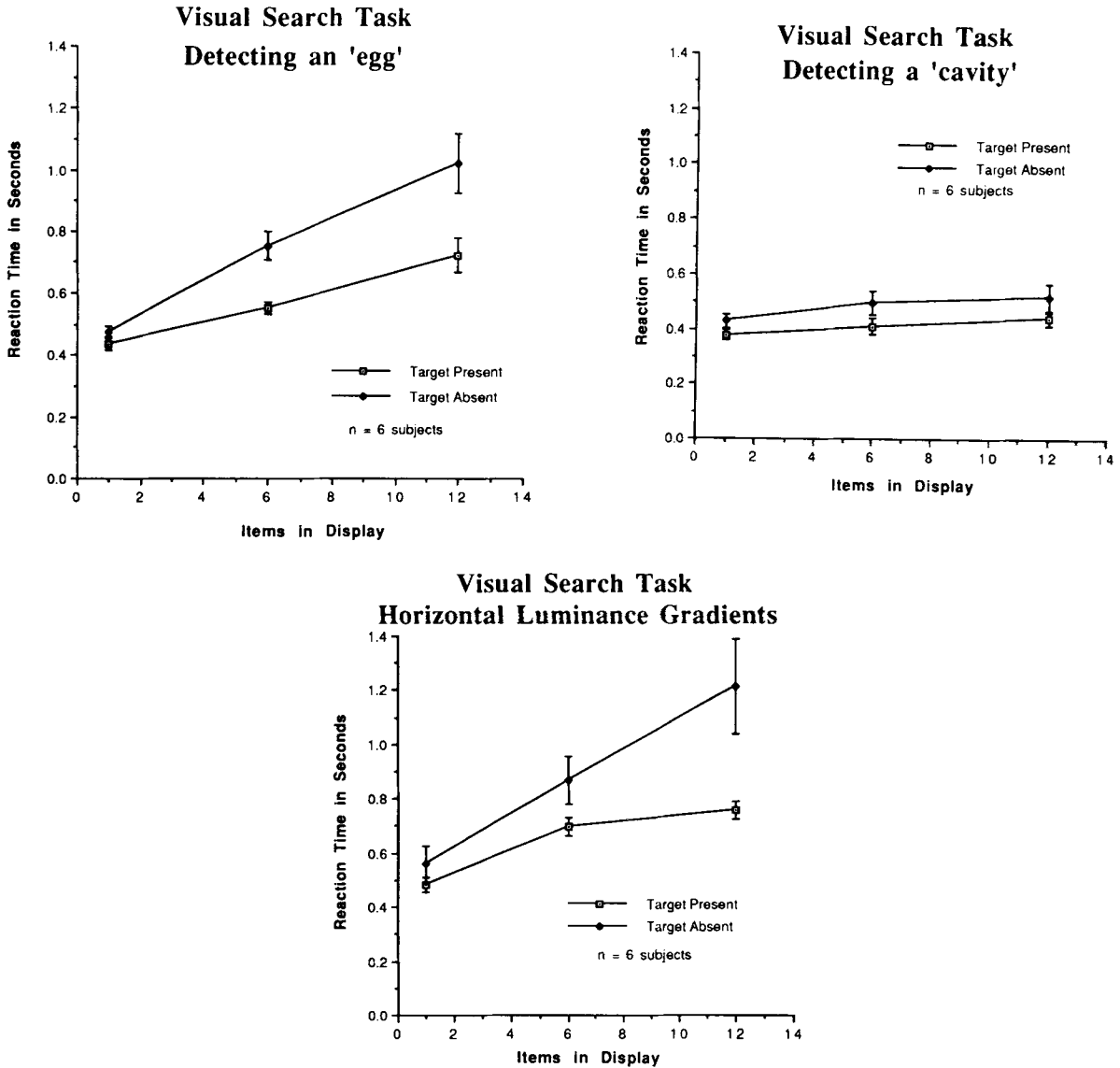
Figure 9. Visual search asymmetries in the extraction of shape from shading. Six naive subjects participated (see text). The reaction time for detecting a "cavity" was unaffected by the number of items in the display. On the other hand, for detecting an "egg," reaction time increased with the number of items in the display and the same was true for detecting left/right shading. These results demonstrate a striking asymmetry in the subjects' ability to detect cavities as opposed to eggs. This effect is seen only in naive subjects. In subjects who have had considerable previous experience with such tasks (as have the authors), the asymmetries do not exist (Kleffner & Ramachandran, 1989).

ANOVA to account for variance. The interpretation of this factor across experimental conditions is ambiguous, but it is included here for completeness. In the first ANOVA, this factor was significant at the .05 level [$F(1,30) = 12.649, p < .0013$], while the interaction between the two factors was not significant. In order to compare the experimental conditions directly, ANOVAs were performed on the data from pairs of experimental conditions. The ANOVA comparing top/bottom shading with bottom/top shading showed that these experimental con-

ditions were significantly different at the .0001 level [$F(1,20) = 48.325, p < .0001$]. The target present/absent factor was significant at the .01 level [$F(1,20) = 8.368, p < .009$]; the interaction was not significant. The ANOVA comparing bottom/top shading and left/right shading was again significant at the .0001 level [$F(1,20) = 22.295, p < .0001$]. (Both the target present/absent factor and the interaction were not significant.) In the ANOVA comparing top/bottom shading with left/right shading, the main effect for target type was not signifi-

cant. The target present/absent factor was significant at the .01 level [$F(1,20) = 12.173, p < .0023$], and the interaction was not significant.

Discussion

These results imply that naive subjects find cavities easier to detect than eggs. This seems surprising and counterintuitive, given the more widespread prevalence of "convexity" in nature (Deutsch & Ramachandran, 1990; Hoffman, 1983), but since virtually nothing is known about the neural detectors that encode shape from shading, we should perhaps be prepared for such surprises.

Treisman and Gormican (1988) argued that search asymmetries arise because the presence of a feature is easier to detect than its absence. For example, a purple object is easy to detect against a background of red objects, because the purple has an "extra" feature—blue—in it and therefore deviates from the "standard" (i.e., red); but a red object cannot be detected as easily against an array of purple objects, since its detection requires the visual system to sense the *absence* of blue. If we accept this logic, we should have to argue that convex objects are the "standard" expected units for the visual system and that cavities are encoded as the same object, but with an extra feature (depth reversal?). This would explain why cavities are easier to detect than eggs.

EXPERIMENT 3
Segregation With Shading

The segregation of figure from ground is another criterion that is sometimes used to decide whether a given visual feature is "elementary" or not (Beck, 1966; Julesz, 1971; Treisman, 1985). It is often assumed that the two criteria pop-out and segregation will necessarily yield the same results, but this is not always true. In certain instances, for example, a target may pop out in a search task, yet when several such targets are present, they cannot be grouped and segregated from the background (Plummer & Ramachandran, 1991). We therefore devised a method that would allow us to directly probe the visual system's ability to achieve perceptual grouping by extracting 3-D shape from shading.

Figure 10A depicts one of the stimuli. Note that instead of the stimulus's being randomly arranged as in Figure 5A, the letter O is composed of eggs displayed against a background of cavities. The subjects' task was to simply report whether they saw a complete O or a broken O on any given trial. The position of the "bite" taken out of the O was also varied randomly from trial to trial (Figures 10B–10C).

Pilot experiments suggested that naive subjects often experience considerable initial difficulty with this task, just as they do when trying to detect a complex cyclopean shape in one of Julesz's (1971) random-dot stereograms. We therefore exposed each subject to a "priming" stimulus, which consisted of the letter X depicted by larger scale shape-from-shading tokens (Figure 11),

before we actually began the forced-choice discrimination experiment.

Method

Subjects. Seven subjects participated. They were drawn from the undergraduate subject pool at the University of California, San Diego, and were naive with respect to the purpose of the experiment.

Display. These displays were also generated on a CRT driven by an Amiga microcomputer. Each stimulus consisted of a circle made up of target items surrounded by a field of distractor items as in Figure 10A. The circle was made up of 12 target items, which were arranged loosely in a circle with a radius subtending 4.1°, against a background of 39 distractors. The targets and distractors, the same pairs that were used in Experiment 1, are illustrated in Figure 6. Each target or distractor item subtended .6°. On half of the trials, a "broken" circle was constructed by replacing 3 consecutive targets in the circle with distractors. The position of the break in the circle was selected randomly from the 12 possible positions. Before the test stimulus was presented, the subjects were shown a preexposure stimulus, which consisted of an X pattern (shown in Figure 11) composed of targets and distractors that were 1.0° across. The targets and distractors, constructed from 16 luminance levels from .057 to 136.1 cd/m², were presented on a background of 14.6 cd/m².

Procedure. The subjects were seated .75 m from the screen in a dark room. Each trial began with the presentation of the preexposure stimulus for 4.0 sec, followed by a dark screen for .1 sec. The test stimulus was then presented for 1.1 sec, after which time the screen went dark. The subjects' task was to determine whether the circle was complete or broken. The subjects were allowed to respond at any time during or after the stimulus presentation, using two keys on the keyboard. The subjects' responses and reaction times were recorded. The subjects were given two training blocks of 80 randomly mixed trials, followed by the experimental block (80 trials, randomly mixed).

Results

The percent correct performance from the experiment is shown in Figure 12. A one-way ANOVA with repeated measures was performed; all experimental conditions were included, with percent correct as the dependent variable. The main effect for type of shading was significant at the .05 level [$F(2,12) = 5.69, p < .018$]. Thus, the degree of segregation obtained varied significantly with the type of shading in the targets and distractors. A separate ANOVA with only two experimental conditions—top shading and side shading—also produced a significant main effect for type of shading at the .05 level [$F(1,6) = 10.179, p < .019$]. A similar comparison of top shading with a step change in luminance was again significant at the .05 level [$F(1,6) = 9.66, p < .021$]. The difference between side shading and a step change in luminance was not significant.

Interestingly, the subjects' reaction times also varied with the type of shading, even though the experimental displays were on for a brief, fixed period of time. The subjects' responses (see Figure 13) were fastest when they were also the most accurate. These differences in reaction time were significantly different at the .01 level [$F(2,12) = 7.059, p < .009$]. A separate comparison of the reaction times for top shading and side shading produced a significant main effect at the .01 level [$F(1,6) = 15.226, p < .008$].
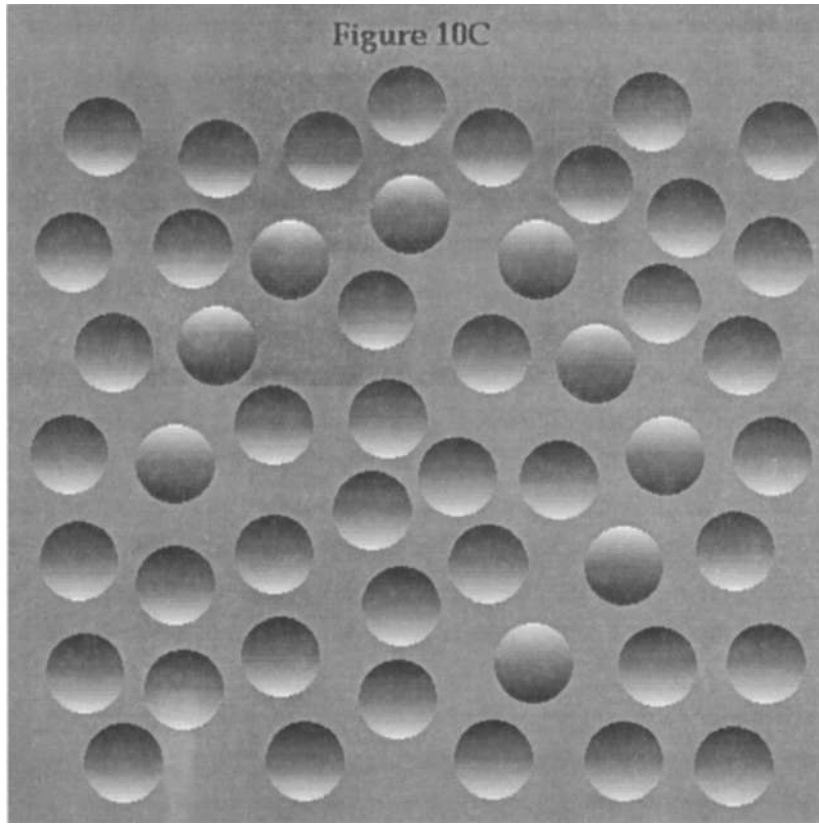
Figure 10A
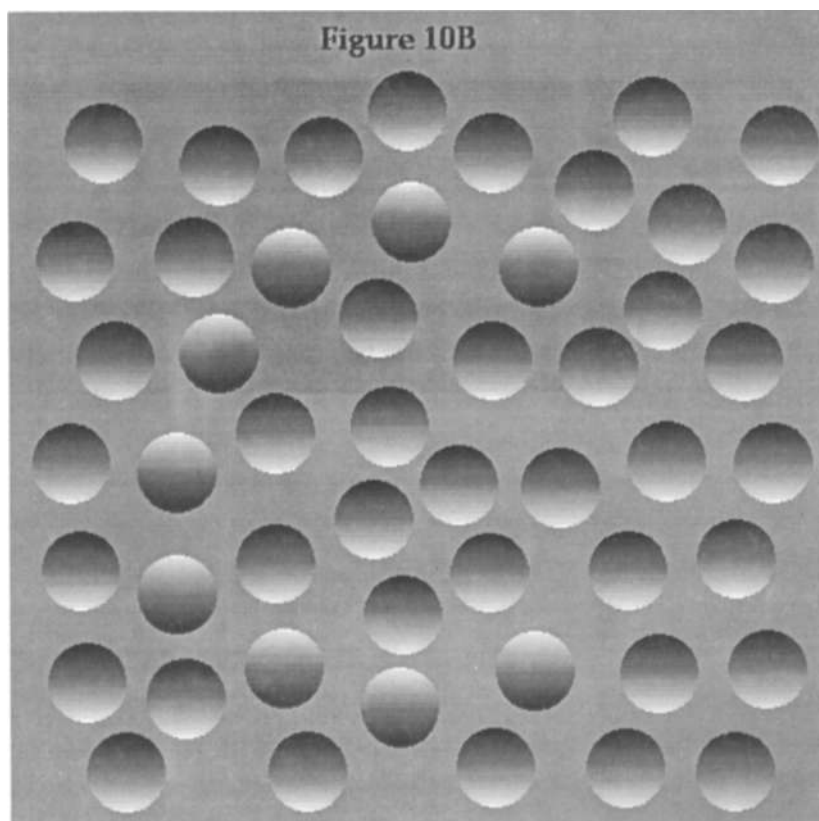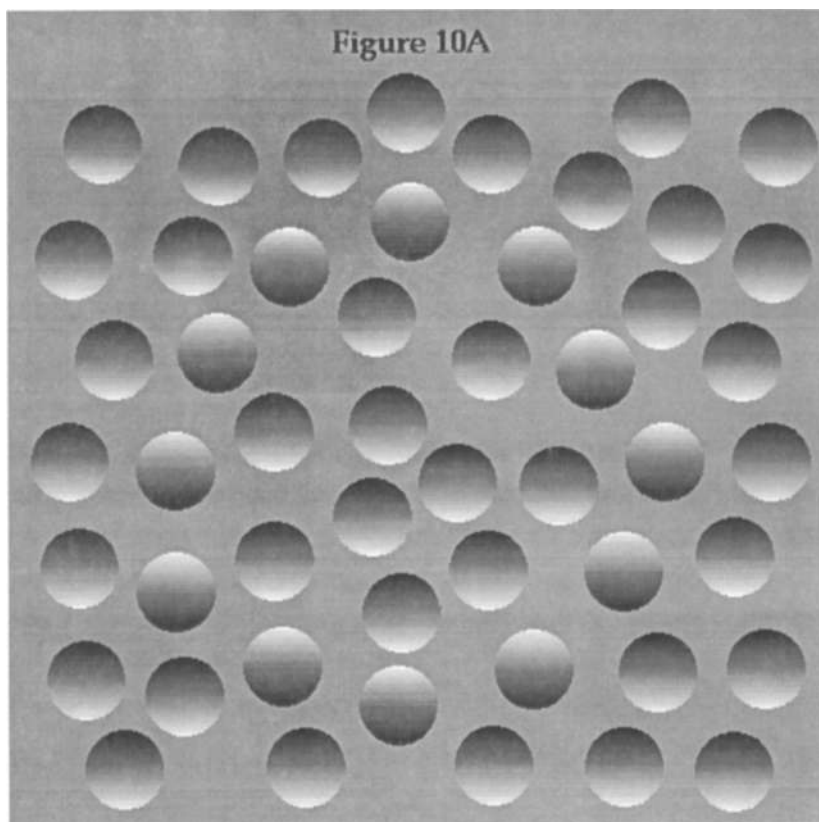
Figure 10B

Figure 10. (Opposite page and above). Sample stimulus used to investigate figure and ground segregation (A). A circle was constructed from the target items (eggs) and presented against a background of distractors (cavities). In half of the trials (B and C), the circle was incomplete (three consecutive eggs missing), and the subjects' task was to determine whether the circle was complete or incomplete within a fixed presentation time.

The main effect for reaction times when top shading was compared with a step change in luminance was not significant [$F(1,6) = 4.458, p < .079$].

Discussion

This experiment shows that the extraction of shape from shading can provide a basis for perceptual grouping when the direction of shading is from top to bottom, but not when the shading is from left to right. The fact that a step change in luminance (as opposed to a continuous gradient) is also relatively ineffective supports our contention that the grouping is based on differences in shading, not on differences in luminance polarity.

These observations suggest that the visual system acts as though it assumes that the sun is shining from above. But how does the visual system know "above" from "below"? Is it the object's orientation in relation to the retina that matters, or its orientation with respect to gravity? This question was originally raised by Yonas, Kuskowski, and Sternfels (1979), who performed a series of ingenious experiments in which they presented ambiguous stimuli to 3-year-old and 7-year-old children. (They used photographs of real objects rather than computer-generated images.) They found that the responses of the 3-year-olds de-

pended almost exclusively on retinal orientation, whereas 7-year-olds showed roughly equal dependence on both retinal and gravitational frames of reference. These results suggest that as children grow older, they progressively shift their responses toward more abstract frames of reference. Curiously, Yonas et al. (1979) did not test adults, but their results imply that if the same trend continues, adults should show a still higher dependence on gravitational (rather than retinal) "upright."

Ramachandran (1988a, 1988b) tested this hypothesis by presenting stimuli such as Figure 2 to adult subjects and asking them to rotate their heads by 90°. Instantly, all the objects that were "top" lit *in relation to the retina* were seen as convex, and the others were seen as cavities. The effect was striking; even a head tilt as little as 15°–20° was sufficient to generate the unambiguous percept of eggs and cavities. When the head was tilted by 15° in the opposite direction, the eggs and cavities reversed depth instantly. We recently showed a slide of this display to a lay audience of several thousand spectators (Ramachandran, 1989a), and most of them reported the perceptual switch. We may conclude from this that the interpretation of shape from shading depends primarily, if not exclusively, on retinal rather than gravitational cues.
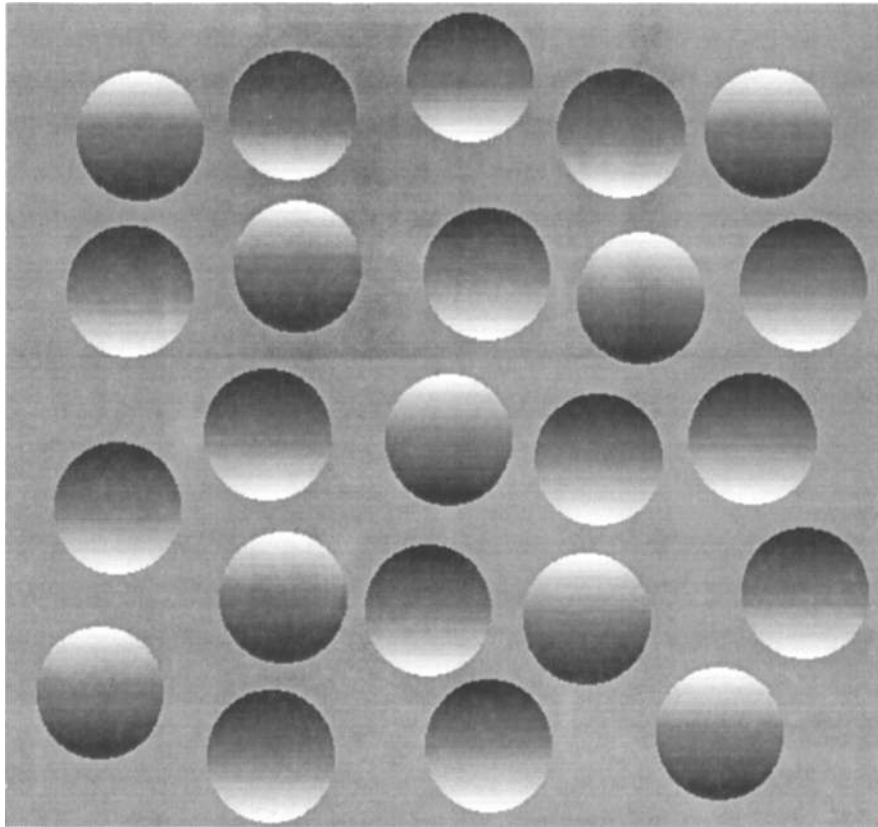
Figure 11. This is an example of the preexposure stimulus presented before each experimental trial. The subjects viewed a display containing an X shape (composed of target items) for 4.0 sec before the experimental stimulus was presented.
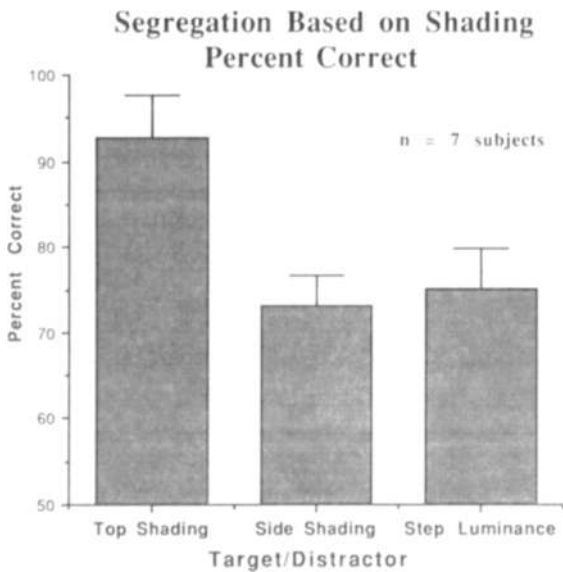

Figure 12. Results from the basic segregation task with shading (Experiment 3). The subjects were much more accurate at determining whether the circle was complete or incomplete with vertical shading than they were with either horizontal shading or a step change in luminance.

The implication is that shape is probably extracted from shading fairly early in visual processing, since it is not subject to vestibular correction for head tilt. The reason for the slight discrepancy with the results of Yonas et al. (1979) is unclear. One possibility is that the verbal reports of 7-year-olds are inherently unreliable. A second, more likely, possibility is that the test we used (perceptual segregation of eggs from cavities) is a more "objective"—and perhaps more sensitive—measure of the extraction of shape from shading than is simply judging convexity versus concavity. Finally, the fact that Yonas et al. used photographs of real objects rather than computer-generated displays might have contributed to the differences in results.

To test the retinocentric hypothesis more formally, we used the segregation task developed for Experiment 2. In half of the trials, the subjects sat upright in front of the CRT screen; in half, they lay down on their sides so that they viewed the CRT screen from a 90° angle.

EXPERIMENT 4
Retinal Versus Gravitational Coordinates

Method

Subjects. The subjects were 11 undergraduates from the subject pool at the University of California, San Diego.
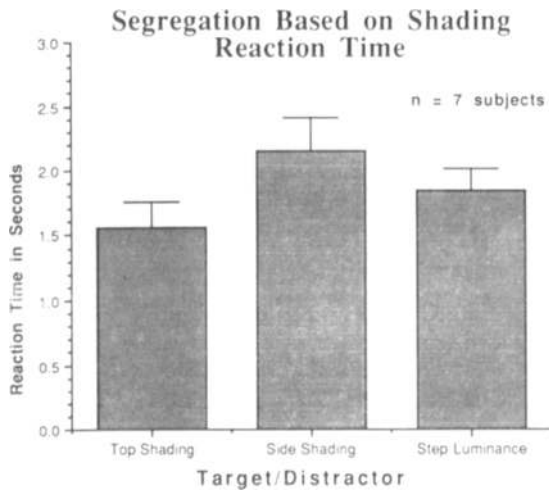
Segregation Based on Shading Reaction Time

Figure 13. The subjects' reaction times for the segregation task (Experiment 3) varied with the type of shading used to depict the C or O shape, even though total duration of presentation was constant for all displays (1.1 sec).

Display. The displays were identical to those in Experiment 1 (see Figure 6), with the exception that the step change in luminance was not used. The two types of shaded stimuli (top and side) were compared from two viewing conditions.

Procedure. The procedure was identical to that in Experiment 3, except that the subjects sat with the head upright during alternate experimental blocks and lay down directly in front of the screen with the head at a 90° angle to the screen in the other half. In both cases, the subjects' eyes were .75 m from the screen.

Results

The results are shown in Figure 14. A two-way ANOVA comparing both direction of shading and viewing position was not significant for either term, since the results are in opposite directions for the two factors. The interaction was significant at the .01 level [$F(1,10) = 14.492, p < .003$]. This means that the direction of shading that produced segregation varied with the subjects' head position. A one-way ANOVA comparing top shading with side shading for each of the viewing positions was performed, and both comparisons were significant at the .05 level [$F(1,10) = 7.612, p < .020$, for the upright condition, and $F(1,10) = 8.563, p < .015$, when subjects were lying on their sides].

Once again, subjects' reaction times varied with the type of shading and the viewing position, even though the experimental displays were on for a fixed period of time—providing further evidence for the difference in segregation. These data are presented in Figure 15. As with the segregation data, a two-way ANOVA showed a significant interaction at the .01 level [$F(1,10) = 13.720, p < .004$]. A one-way ANOVA comparing top and side shading while the subjects were sitting up was not significant [$F(1,10) = 1.355, p < .271$]. A comparison of top and side shading while subjects were reclining was highly significant at the .001 level [$F(1,10) = 27.256, p < .0004$].

Discussion

These results indicate that the interpretation of shape from shading depends primarily, if not exclusively, on retinal rather than gravitational cues. Obviously, since we often do tilt our heads inadvertently, it would be more

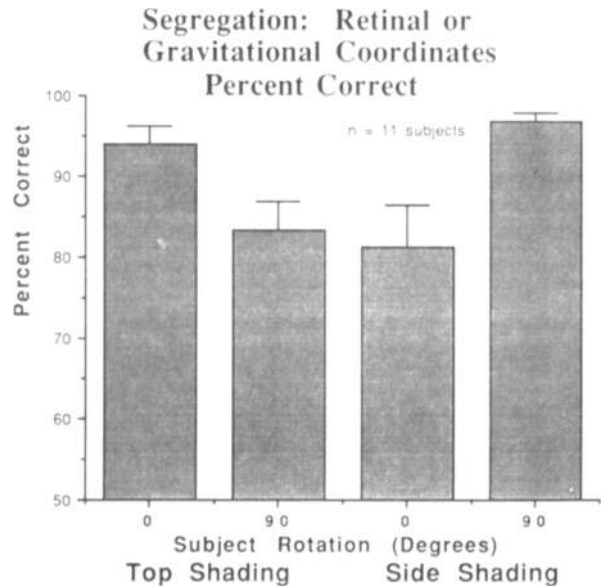## Segregation: Retinal or Gravitational Coordinates Percent Correct

Figure 14. Are shading effects tied to retinal or gravitational coordinates? We explored this by having subjects view the screen while they were either sitting up or lying down. The results indicate that the effect depends primarily, if not exclusively, on retinal rather than gravitational cues (Experiment 4).

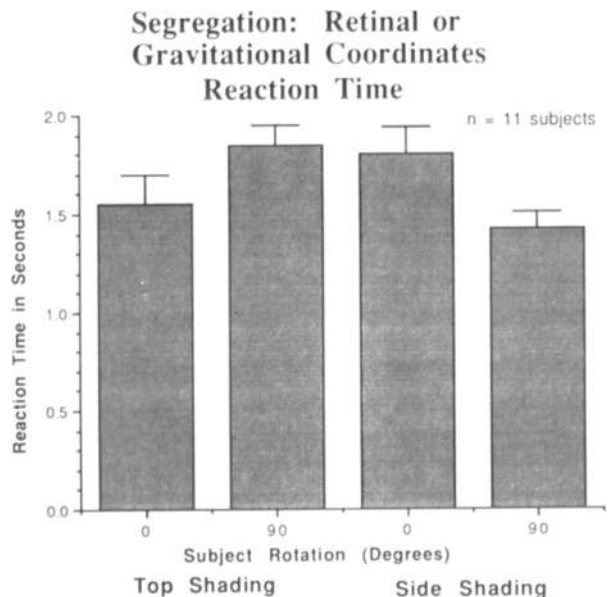## Segregation: Retinal or Gravitational Coordinates Reaction Time

Figure 15. The subjects' reaction times provide evidence that the effect depends on retinal rather than gravitational cues. The subjects responded more quickly when the direction of shading was vertical in *retinal* coordinates, regardless of whether they were sitting up or prone (Experiment 4).

sensible for the visual system to use world-centered co-ordinates, but our results suggest that this is not the case. The curious implication of this is that the mechanisms that compute shape from shading do not correct for head tilt and blindly assume that the sun is stuck to the head when one tilts one's head (or body) by 90°!

Finally, if the extraction of 3-D shape from shading is carried out fairly early in visual processing, as we have suggested, one would also expect it to interact with other "front end" visual mechanisms such as motion perception. Our next experiment was designed to explore this possibility.

EXPERIMENT 5
Shape From Shading As an Input to Motion Perception

Can shapes that are conveyed exclusively through differences in shading (e.g., the O in Figure 10A) be used by the visual system to establish motion correspondence? To find out, we constructed a three-frame apparent motion sequence in which the figure defined by shading was displaced in either the left or right direction along the x-axis. The actual positions of all the eggs and cavities were uncorrelated in successive frames, so that correspon-

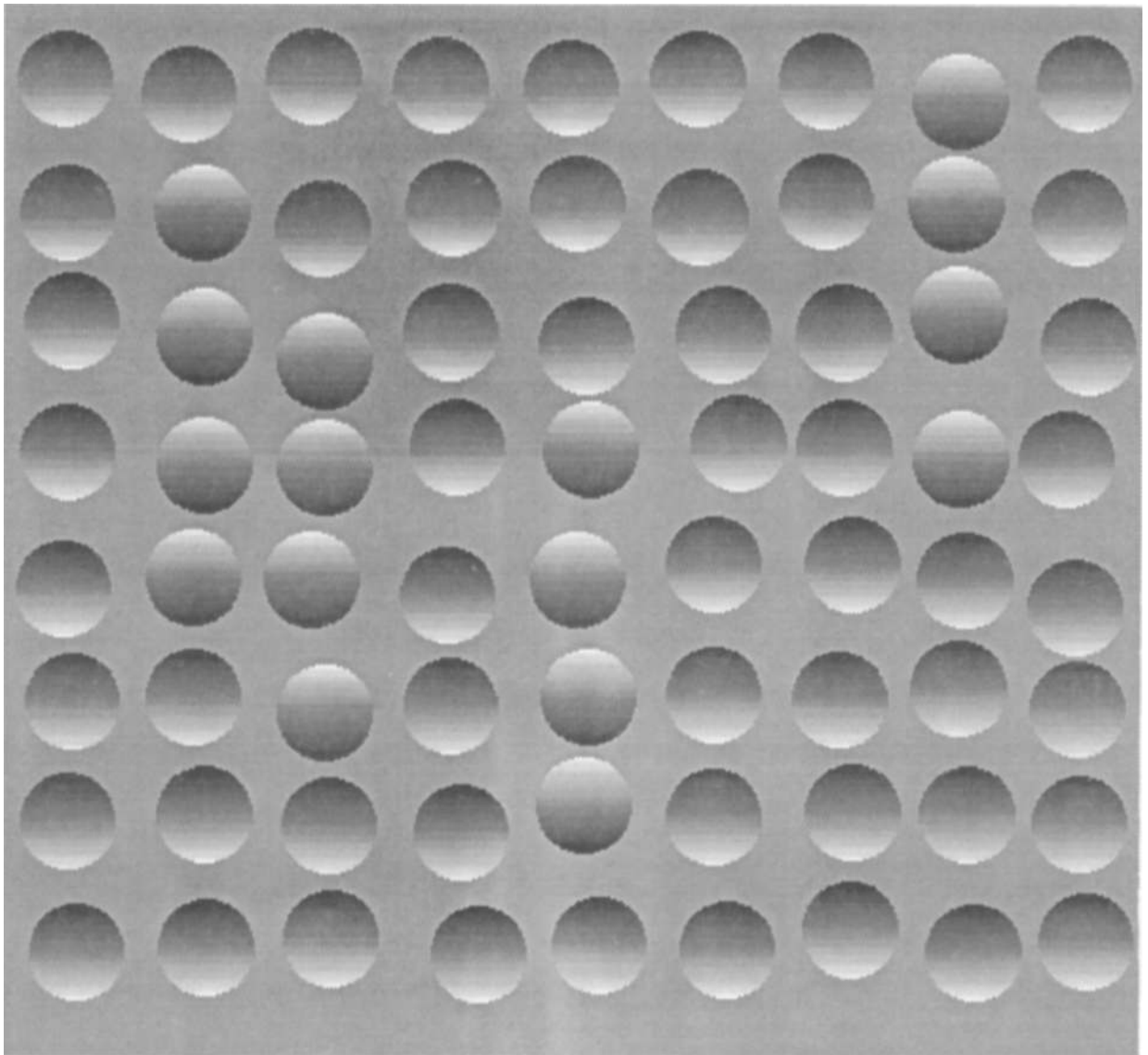Figure 16. Sample stimulus used in the motion detection task (Experiment 5). The subjects were shown three frames of an apparent motion sequence. Each frame portrayed a figure defined exclusively by differences in shading. The figure was displaced either to the left or to the right in successive frames. The frames were completely uncorrelated in the luminance domain. The task was to report the direction of motion (left or right).

dence could not be established on the basis of luminance cues. The question is, can subjects use the shading information to report the direction of motion correctly?

Method

Subjects. The subjects were 9 undergraduates from the subject pool at the University of California, San Diego.

Display. These displays were also generated on a CRT driven by an Amiga microcomputer, and the targets and distractors had the same size and construction as did those which were used in Experiment 4 to compare vertical shading with horizontal shading. The stimuli consisted of three frames presented as an apparent motion sequence. Each frame (see Figure 16 for an example) consisted of distractor items presented on a regular grid, 8 rows × 10 columns, in a display 6.3° high × 8.6° wide. The position of each item was varied by a random distance (±.2°) in both the x and y directions, so that the three frames were uncorrelated. The target pattern consisted of four vertical columns (four eggs per column) at random positions in the first frame. The position of the target pattern was moved by one position to the left or right in each of the following frames to provide the basis for apparent motion. Prior to the presentation of the test stimuli, the subjects were shown the preexposure stimulus that had been used in Experiments 3 and 4.

Procedure. The subjects were seated .75 m from the screen in a dark room. Each trial began with the presentation of the preexposure stimulus for 6.0 sec. Each of the three frames of the experimental stimulus was then presented for .4 sec, after which time the screen went dark. The subjects' task was to determine whether the direction of motion was to the left or to the right. The subjects were allowed to respond at any time during or after the stimulus presentation by using two keys on the keyboard. The responses and reaction times were recorded. The subjects were given instructions but no training with the experimental stimuli. They participated in three blocks of 50 trials for each set of targets and distractors.

Results

The results of the experiment are shown in Figure 17. The subjects were able to predict the direction of motion much more accurately when the direction of shading was vertical (top to bottom) rather than horizontal (left to right shading). A one-way ANOVA of percent correct as a function of the target/distractor type was highly significant [at the .001 level; $F(1,8) = 32.985, p < .0004$].

Discussion

These results show clearly that the extraction of shape from shading can contribute to motion processing (Ramachandran, 1988a, 1988b). Again, in order to achieve this, the visual system must "assume" overhead lighting, since motion discrimination is reduced considerably for targets conveyed by left–right gradients. The findings are somewhat surprising, since they imply that even a monocular depth cue such as shading—which is often regarded as "cognitive"—can drive the motion system, which is usually regarded as an early or "front end" visual process. One wonders whether other monocular depth cues such as perspective can also drive the motion system.

Also, our observation that the interpretation of shape from shading—based on the "overhead lighting" assumption—can drive both perceptual segregation and motion perception is inconsistent with the view recently expressed by Reichel and Todd (1990) that this assumption is "only of marginal significance" in natural vision. Of course, as Reichel and Todd point out, the assumption can be overridden by other conflicting cues in the image (such as "height in field" or occlusion); but then, so can any other assumption in perception. For example, even stereopsis can be overridden easily in a hollow mask that is lit from below (Ramachandran, 1988a); yet one would not want to conclude from this that stereopsis is only of marginal significance in natural vision!

GENERAL DISCUSSION

Although the study of shape from shading has attracted considerable attention from the artificial intelligence community (e.g., Bulthoff & Mallot, 1988; Horn, 1975; Lehky & Sejnowski, 1988), there have been very few psychophysical studies of how shading is extracted by the human visual system.

Taken collectively, our results suggest that the extraction of shape from shading incorporates two assumptions—that there is only one light source illuminating most of the image, and that the light is shining from above. These assumptions seem to affect not only the sign of perceived depth (i.e., convex vs. concave, as seen in Figures 2 and 5A) but also its magnitude (Figure 4). In addition to these two constraints, there appears to be a weaker "default" assumption that objects are more likely to be convex rather than concave (Deutsch & Ramachandran, 1990). This would explain why naive subjects usually see the objects

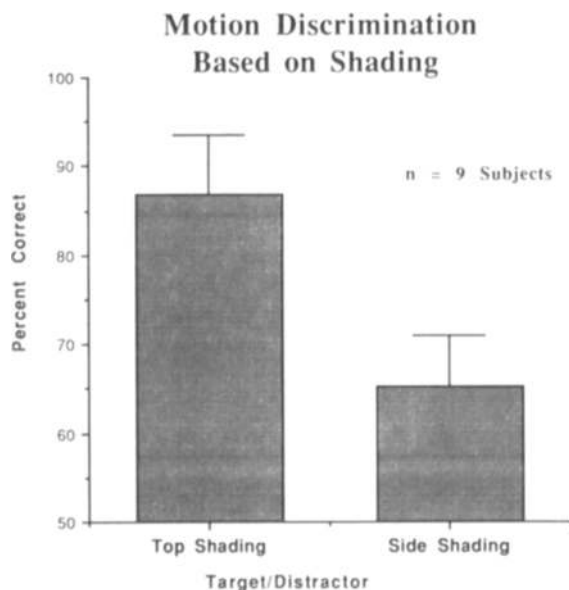**Motion Discrimination Based on Shading**

n = 9 Subjects

Figure 17. Results from the motion detection task (Experiment 5). Subjects were more accurate at determining the direction of motion with vertical shading than they were with either horizontal shading or a step change in luminance.

in Figure 1 as convex, and why it takes some effort to see them as concave.

Furthermore, we find that once these 3-D shapes have been extracted, they can serve as a basis for pop-out and for perceptual grouping. Since these effects are observed only for top-bottom differences in shading, our results imply that relatively complex "scene-based" image characteristics such as direction of lighting can influence visual search and figure-ground segregation (Ramachandran, 1988a, 1988b; see also Enns & Rensink, 1990).

The conclusion that more complex "whole image" characteristics can influence perceptual grouping also receives support from another experiment that we recently carried out to study motion perception (Plummer & Ramachandran, 1991; Ramachandran & Rogers-Ramachandran, 1991). We began with two sparse patterns (A and B) that were optically superimposed on each other. Each pattern was composed of randomly arranged small circles. We then made one of the patterns (A) approach the observer so that the circles moved radially outward from the center, while, at the same time, the other was made to shrink inward (i.e., to "recede" from the observer). The sizes of the circles were randomized, and we also presented the whole display through a window so that the outer margins of A and B were invisible. We found that the subjects had no difficulty in segregating A from B so that what they saw was a pattern receding *through* an approaching plane of circles. Notice that in each plane (A or B), there were elements that were actually moving in *opposite* directions in the frontoparallel plane—corresponding to either expansion or contraction—yet the visual system had no difficulty in grouping these together. We suggest, therefore, that although *segregation* is usually based on local feature differences, grouping can take advantage of more "global" rules that reflect higher order invariances. As a control, we used a very similar display in which all the individual circles of a pattern were made to expand, but there was no global expansion of the pattern as a whole (i.e., the distances between the centers of the circles did not change). The circles in Pattern B were made to shrink simultaneously. No grouping or segregation was observed in this display.

Our third experiment showed that the overhead light assumption is based on retinal rather than phenomenal or world-centered coordinates. This finding also suggests that the extraction of shape from shading is unlikely to be very cognitive and that it is extracted fairly early in visual processing—certainly earlier than vestibular and cognitive correction for head tilt. This finding is surprising, since it implies that, at least as far as the extraction of shading is concerned, the visual system assumes that the sun moves with the head! If the visual system is indeed "intelligent" as some have argued, why does it incorporate such a primitive assumption? One possibility is that even though we do tilt our heads occasionally, *statistically speaking* we do walk upright most of the time, and so, the visual system can get away with this primitive assumption. The advantage, of course, is that extraction of

shape from shading can then proceed much more quickly without the additional computational burden of having to correct for head tilt—a process that might be very time consuming. This line of reasoning accords well with our view (Ramachandran, 1985, 1990) that perception often involves the use of "short-cuts"—heuristics, rather than sophisticated, optimally designed algorithms.

The importance of overhead lighting as a "natural constraint" is also consistent with the observation that many plains-dwelling animals (e.g., gazelles, cheetahs, etc.) have evolved "countershading"; that is, they have pale bellies that serve to neutralize the shading produced by the sun shining from above. Our results suggest that countershading may be effective mainly because it reduces the extent to which an animal's shape pops out from the background. Curiously, there is a species of caterpillar that displays *reverse* countershading (i.e., a dark belly instead of a pale belly)—an observation that does not make sense unless one realizes that this species habitually hangs upside down from twigs (Tinbergen, 1968)! And finally, it has been shown recently (Greenwood, 1991) that certain octopuses can actually reverse their shading in a matter of seconds if deliberately held upside down—a "shading reflex" that is thought to be vestibular rather than visual in origin.

Is the segregation in Figure 5A due to perceived depth or is it due to 3-D shape? Note that the front surfaces of the eggs are, on the whole, nearer than the margins (or inner surfaces) of the cavities; perhaps this difference in depth leads to the grouping and segregation observed in Figure 5A. To explore this, we tried presenting the eggs and cavities in Figure 5A in random stereoscopic planes, so that some of the cavities were actually stereoscopically nearer than the eggs. When we viewed this display, we found, to our surprise, that the eggs could still be grouped effortlessly and segregated from the cavities, even though they occupied random depth planes. We concluded, therefore, that the segregation observed in these displays is based on 3-D *shape* (or perhaps even directly on the shading), rather than on the perceived depth (see also Ramachandran, 1990).

Another interesting effect that we have recently observed is that of background luminance on perceptual grouping and pop-out. We found that segregation was optimal when we used a neutral gray background whose luminance was identical to the mean luminance of the shaded tokens (Kleffner & Ramachandran, 1989). When the background was too light or too dark (e.g., see Figure 18), the degree of segregation was reduced considerably. The observation suggests that the visual system tends to "assume" that the background has the same reflectance characteristics as do the objects in the foreground (i.e., that it is made of the same material as they are). This may seem surprising, since the assumption that the background and the objects lying on it share the same material is not generally true for most objects. It is certainly true for lumpy terrain, however. Could the shape from shading system have evolved primarily as a primitive visual mod-

Visual Search Task
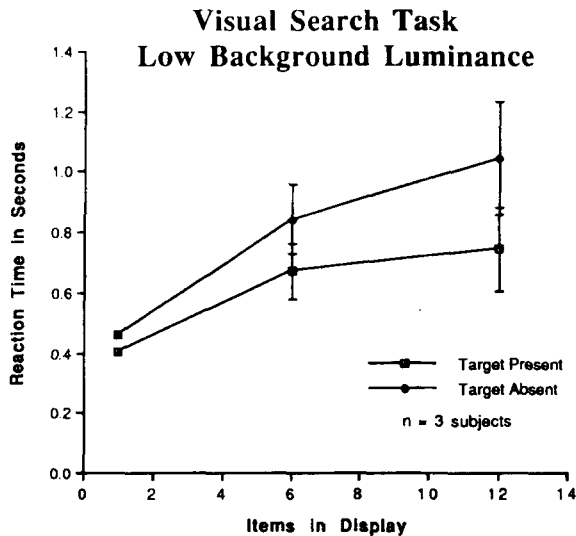## Low Background Luminance

Figure 18. Results from a visual search task in which the background luminance (.057 cd/m²) was lower than the mean luminance of the target (an egg) and distractors (cavities). Results indicate that 3 subjects' ability to find the target was reduced.

ule to control locomotion and stop us from falling over bumps and hollows?

The influence of "top-down" effects on the perception of shape from shading also deserves further study. It is known, for example, that a hollow mask tends to look "normal" (i.e., convex) even if this requires vetoing both stereo disparity cues (Gregory, 1970; Helmholtz, 1910) and the assumption of overhead lighting (Ramachandran, 1988a, 1988b). It is unclear, however, whether the effect derives from familiarity with faces or from a more general "convexity" assumption (Deutsch & Ramachandran, 1990). In a recent experiment, we tried viewing the inside of a hollow mask with the nose turned inside out so that it conveyed "normal" crossed disparities. Interestingly, we found that although the depth of the entire face reversed, the nose continued to look convex so that the net result was a completely normal face. Hence the reversal-of-relief effect is not a global operation; it does not simply involve changing all the " + " signs to " − " signs and vice versa (Ramachandran & Gregory, 1991). On the contrary, our results imply that the reversal of depth is applied by the visual system only to the parts of the object where it is deemed necessary. The effect was especially convincing when we moved our heads, since the face appeared to "follow" us, but the nose appeared to move in the opposite direction!

Experiment 5 shows that the extraction of shape from shading can contribute to motion perception and that it may give us some hint about the neural locus at which shading is extracted by the visual system. The physiology of motion perception has been studied extensively; we know that much of the processing occurs in layer 4B in area 17 and continues into the broad stripes of area 18

and in the middle temporal areas (MT) in the parietal lobes (Allman, 1987; Livingstone & Hubel, 1987; Van Essen, 1979; Zeki, 1978). It seems reasonable to assume, therefore, that the extraction of shape from shading must occur either within one of these areas or at an earlier stage. Just as the introduction of random-dot stereograms by Julesz (1971) prompted a search for disparity-detecting neurons in the visual cortex (e.g., see Pettigrew, 1972), one hopes that our findings will motivate physiologists to look for cells in one of these areas—in MT or MST, for example—that extract shading information. One way to begin such a search would be to confront motion sensitive cells in MT with moving targets similar to those in Figure 16. Would such cells also respond to motion that is conveyed through shape from shading?

REFERENCES

ALLMAN, J. (1987). Evolution of the brain in primates. In R. L. Gregory, (Ed.), *Oxford companion to the mind* (pp. 633-639). Oxford University Press.

BECK, J. (1966). Effect of orientation and of shape similarity on perceptual grouping. *Perception & Psychophysics*, 1, 300-302.

BERBAUM, K., BEVER, T., & CHUNG, C. (1983). Light source position in the perception of object shape. *Perception*, 12, 411-415.

BERGSTROM, S. S. (1987). *Color constancy: Support to a vector model for the perception of illumination, colour, and depth* (Applied Psychology Rep. No. 26). Umeå, Sweden: University of Umeå, Department of Psychology.

BREWSTER, D. (1847). On the conversion of relief by inverted vision. *Edinburgh Philosophical Transactions*, 15, 657.

BULTHOFF, H. H., & MALLOT, H. A. (1988). Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America*, 5, 1749-1758.

DEUTSCH, J. A., & RAMACHANDRAN, V. S. (1990). Binocular depth reversals despite familiarity cues: An artifact? *Science*, 249, 565-566.

ENNS, J. T., & RENSINK, R. A. (1990). Influence of scene-based properties on visual search. *Science*, 247, 721-723.

GREENWOOD, J. (1991). Marine quick-change acts. *Nature*, 349, 741-742.

GREGORY, R. L. (1970). *The intelligent eye*. New York: McGraw-Hill.

HELMHOLTZ, H. L. F. VON (1910). *Handbuch der Physiologischen Optik*. Hamburg und Leipzig: Leopold Voss.

HOFFMAN, D. D. (1983). The interpretation of visual illusions. *Scientific American*, 249(6), 154-162.

HORN, B. (1975). Obtaining shape from shading information. In P. H. Winston (Ed.), *Psychology of computer vision*. New York: McGraw-Hill.

HOWARD, I. P. (1983). Occluding edges in apparent reversal of convexity and concavity. *Perception*, 12, 85-86.

JULESZ, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.

KLEFFNER, D., & RAMACHANDRAN, V. R. (1989). Perception of three-dimensional shape-from-shading. *Investigative Ophthalmology & Visual Science*, 30(Suppl.), 252.

LAND, E. H. (1983). Recent advances in retinex theory and some implications for cortical computations: Color vision and the natural image. *Proceedings of the National Academy of Sciences*, 80, 5163-5169.

LEHKY, S. R., & SEJNOWSKI, T. J. (1988). Network models of shape-from-shading: Neural function arises from both receptive and projective fields. *Nature*, 333, 452-454.

LIVINGSTONE, M., & HUBEL, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement and depth. *Journal of Neuroscience*, 7, 3416-3468.

NAKAYAMA, K., & LOOMIS, J. M. (1974). Optical velocity patterns; velocity sensitive neurons and space perception: A hypothesis. *Perception*, 3, 63-80.

PETTIGREW, D. (1972). The neurophysiology of binocular vision. *Scientific American*, **227**, 84.

PLUMMER, D. J., & RAMACHANDRAN, V. S. (1991). *Perceptual grouping based on expansion/contraction and motion capture*. Paper presented at Conference on Neural Networks for Vision and Image Processing, Boston University.

RAMACHANDRAN, V. S. (1985). The neurobiology of perception: Guest editorial. *Perception*, **47**, 97-105.

RAMACHANDRAN, V. S. (1988a). Perceiving shape from shading. *Scientific American*, **269**, 76-83.

RAMACHANDRAN, V. S. (1988b). Perception of shape from shading. *Nature*, **331**, 163-166.

RAMACHANDRAN, V. S. (1989a, November). Presidential lecture given at the meeting of the Society for Neurosciences, Phoenix, AZ.

RAMACHANDRAN, V. S. (1989b). *2-D or not 2-D; that is the question*. Lecture given at Symposium on Vision, Art & Brain, Bristol, England. To appear in R. L. Gregory, J. Harris, & P. Heard (Eds.), *The artful brain*. Cambridge: Cambridge University Press.

RAMACHANDRAN, V. S. (1990). Visual perception in people and machines. In A. Blake & T. Troscianko (Eds.), *AI and the eye*. Bristol, England: Wiley.

RAMACHANDRAN, V. S., & GREGORY, R. L. (1991). *Unmasking the truth*. Manuscript in preparation.

RAMACHANDRAN, V. S., & ROGERS-RAMACHANDRAN, D. C. (1991). Phantom contours: A new class of visual patterns that selectively activates the magnocellular pathway in man. *Bulletin of the Psychonomic Society*, **29**, 391-394.

REICHEL, F. D., & TODD, J. T. (1990). Perceived depth inversion of smoothly curved surfaces due to image orientation. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 653-664.

RITTENHOUSE, D. (1786). Explanation of an optical deception. *Transactions of the American Philosophical Society*, **2**, 37-42.

THAYER, G. H. (1909). *Concealing coloration in the animal kingdom*. New York: MacMillan.

TINBERGEN, N. (1968). *Curious naturalists*. New York: Doubleday.

TODD, J. T., & MINGOLLA, E. (1983). Perception of surface curvature and direction of illumination from patterns of shading. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 583-595.

TREISMAN, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics, & Image Processing*, **31**, 156-177.

TREISMAN A. (1986). Features and objects in visual processing. *Scientific American*, **255**, 114-126.

TREISMAN, A., & GORMICAN, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, **95**, 15-48.

VAN ESSEN, D. C. (1979). Visual cortical areas. *Annual Review of Neuroscience*, **2**, 227-263.

YONAS, A., KUSKOWSKI, M., & STERNFELS, S. (1979). The role of frames of reference in the development of responsiveness to shading. *Child Development*, **50**, 495-500.

ZEKI, S. M. (1978). Functional specialization in the visual cortex of the rhesus monkey. *Nature*, **274**, 423-428.

NOTE

1. We are happy to acknowledge that J. T. Enns and R. A. Rensink (personal communication, 1991) have independently observed a similar effect. Note that the asymmetry is seen only in naive subjects (Figure 9A) but not in experienced subjects (Figure 7A). Not surprisingly, extensive practice with our stimuli seems to reduce the asymmetry.